Tight Bounds on 3-Team Manipulations in Randomized Death Match

Atanas Dinev[1(✉)] and S. Matthew Weinberg[2]

[1] Massachusetts Institute of Technology, Cambridge, MA 02142, USA
adinev@mit.edu
[2] Princeton University, Princeton, NJ 08544, USA
smweinberg@princeton.edu

Abstract. Consider a round-robin tournament on n teams, where a winner must be (possibly randomly) selected as a function of the results from the $\binom{n}{2}$ pairwise matches. A tournament rule is said to be k-SNM-α if no set of k teams can ever manipulate the $\binom{k}{2}$ pairwise matches between them to improve the joint probability that one of these k teams wins by more than α. Prior work identifies multiple simple tournament rules that are 2-SNM-1/3 (Randomized Single Elimination Bracket [17], Randomized King of the Hill [18], Randomized Death Match [6]), which is optimal for $k = 2$ among all Condorcet-consistent rules (that is, rules that select an undefeated team with probability 1).

Our main result establishes that Randomized Death Match is 3-SNM-(31/60), which is tight (for Randomized Death Match). This is the first tight analysis of any Condorcet-consistent tournament rule and at least three manipulating teams. Our proof approach is novel in this domain: we explicitly find the most-manipulable tournament, and directly show that no other tournament can be more manipulable.

In addition to our main result, we establish that Randomized Death Match disincentivizes Sybil attacks (where a team enters multiple copies of themselves into the tournament, and arbitrarily manipulates the outcomes of matches between their copies). Specifically, for any tournament, and any team u that is not a Condorcet winner, the probability that u or one of its Sybils wins in Randomized Death Match approaches 0 as the number of Sybils approaches ∞.

1 Introduction

Consider a tournament on n teams competing to win a single prize via $\binom{n}{2}$ pairwise matches. A tournament rule is a (possibly randomized) map from these $\binom{n}{2}$ matches to a single winner. In line with several recent works [1,2,6,17,18], we study rules that satisfy some notion of fairness (that is, "better" teams should be more likely to win), and non-manipulability (that is, teams have no incentive to manipulate the matches).

More specifically, prior work identifies *Condorcet-consistence* (Definition 4) as one desirable property of tournament rules: whenever an undefeated team

K. A. Hansen et al. (Eds.): WINE 2022, LNCS 13778, pp. 273–291, 2022.
https://doi.org/10.1007/978-3-031-22832-2_16

exists, a Condorcet-consistent rule selects that team as the winner with probability 1. Another desirable property is *monotonicity* (Definition 6): no team can unilaterally increase the probability that it wins by throwing a single match. Arguably, any sensible tournament rule should at minimum satisfy these two basic properties, and numerous such simple rules exist.

[1,2] further considered the following type of deviation: what if the same company sponsors multiple teams in an eSports tournament, and wants to maximize the probability that one of them wins the top prize?[1] In principle, these teams might manipulate the outcomes of the matches they play amongst themselves in order to achieve this outcome. Specifically, they call a tournament rule k-Strongly-Non-Manipulable (k-SNM, Definition 5), if no set of k teams can successfully manipulate the $\binom{k}{2}$ pairwise matches amongst themselves to improve the probability that one of these k teams wins the tournament. Unfortunately, even for $k = 2$, [1,2] establish that no tournament rule is both Condorcet-consistent and 2-SNM.

This motivated recent work in [6,17,18] to design tournament rules which are Condorcet-consistent *as non-manipulable as possible*. Specifically, [17] defines a tournament rule to be k-SNM-α if no set of k teams can manipulate the $\binom{k}{2}$ pairwise matches amongst themselves to increase total probability that any of these k teams wins *by more than* α (see Definition 5). These works design several simple Condorcet-consistent and 2-SNM-1/3 tournament rules, which is optimal for $k = 2$ (see [17]). In fact, the state of affairs is now fairly advanced for $k = 2$: each of [6,17,18] proposes a new 2-SNM-1/3 tournament rule. [18] considers a stronger fairness notion that they term Cover-consistent, and [6] considers probabilistic tournaments (see Sect. 1.3 for further discussion).

However, *significantly* less is known for $k > 2$. Indeed, only [18] analyzes manipulability for $k > 2$. They design a rule that is k-SNM-2/3 for all k, but that rule is non-monotone, and it is unknown whether their analysis of that rule is tight. Our main result provides a tight analysis of the manipulability of Randomized Death Match (first defined in [6]) when $k = 3$. We remark that this is: a) the first tight analysis of the manipulability of any Condorcet-consistent tournament rule for $k > 2$, b) the first analysis establishing a monotone tournament rule that is k-SNM-α for any $k > 2$ and $\alpha < 1$, and c) the strongest analysis to-date of any tournament rule (monotone or not) for $k = 3$. We overview our main result in more detail in Sect. 1.1 below.

Beyond our main result, we further consider manipulations through a *Sybil attack* (Definition 9). As a motivating example, imagine that a tournament rule is used as a proxy for a voting rule to select a proposal (voters compare each pair of proposals head-to-head, and this constitutes the pairwise matches input to a tournament rule). A proposer may attempt to manipulate the protocol with a Sybil attack, by submitting numerous nearly-identical clones of the same proposal. This manipulates the original tournament, with a single node u_1 corresponding to the proposal, into a new one with additional nodes $u_2, \dots, u_m$ cor-

[1] Similarly, perhaps there are multiple athletes representing the same country or university in a traditional sports tournament.

responding to the Sybils. Each node $v \notin \{u_1, \ldots, u_m\}$ either beats all the Sybils, or none of them (because the Sybil proposals are essentially identical to the original). The questions then become: Can the proposer profitably manipulate the matches within the Sybils? Is it beneficial for a proposer to submit as many Sybils as possible? We first show that, when participating in Randomized Death Match, the Sybils can't gain anything by manipulating the matches between them. Perhaps more surprisingly, we show that Randomized Death Match is *Asymptotically Strongly Sybil-Proof*: as the number of Sybils approaches ∞, the collective probability that a Sybil wins RDM approaches *zero* (unless the original proposal is a Condorcet winner, in which case the probability that a Sybil wins is equal to 1, for any number of Sybils > 0).

1.1 Our Results

As previously noted, our main result is a tight analysis of the manipulability of Randomized Death Match (RDM) for coalitions of size 3. Randomized Death Match is the following simple rule: pick two uniformly random teams who have not yet been eliminated, and eliminate the loser of their head-to-head match.

Informal Theorem 1 *(See Theorem 6). RDM is 3-SNM-$\frac{31}{60}$. RDM is not 3-SNM-α for $\alpha < \frac{31}{60}$.*

Recall that this is the first tight analysis of any Condorcet-consistent tournament rule for any $k > 2$ and the first analysis establishing a monotone, Condorcet-consistent tournament rule that is k-SNM-α for any $k > 2$, $\alpha < 1$. Recall also that previously the smallest α for which a 3-SNM-α (non-monotone) Condorcet-consistent tournament rule is known is 2/3.

Our second result concerns manipulation by Sybil attacks. A Sybil attack is where one team starts from a base tournament T, and adds some number $m-1$ of clones of their team to create a new tournament T' (they can arbitrarily control the matches within their Sybils, but each Sybil beats exactly the same set of teams as the cloned team) (See Definition 9). We say that a tournament rule r is *Asymptotically Strongly Sybil-Proof* (Definition 10) if for any tournament T and team $u_1 \in T$ that is not a Condorcet winner, the maximum collective probability that a Sybil wins (under r) over all of u_1's Sybil attacks with m Sybils goes to 0 as m goes to infinity. See Sect. 2 for a formal definition.

Informal Theorem 2 *(See Theorem 8). RDM is Asymptotically Strongly Sybil-Proof.*

1.2 Technical Highlight

All prior work establishing that a particular tournament rule is 2-SNM-1/3 follows a similar outline: for any T, cases where manipulating the $\{u, v\}$ match could potentially improve the chances of winning are coupled with two cases where manipulation cannot help. By using such a coupling argument, it is plausible that one can show that RDM is 3-SNM-($\frac{1}{2} + c$) for a small constant c.

However, given that Theorem 6 establishes that RDM is 3-SNM-31/60, it is hard to imagine that this coupling approach will be tractable to obtain the exact answer.

Our approach is instead drastically different: we find a particular 5-team tournament, and a manipulation by 3 teams that gains 31/60, and directly prove that this must be the worst case. We implement our approach using a first-step analysis, thinking of the first match played in RDM on an n-team tournament as producing a distribution over $(n-1)$-team tournaments.

The complete analysis inevitably requires some careful case analysis, but is tractable to execute fully by hand. Although this may no longer be the case for future work that considers larger k or more sophisticated tournament rules, our approach will still be useful to limit the space of potential worst-case examples.

1.3 Related Work

There is a vast literature on tournament rules, both within Social Choice Theory, and within the broad CS community [3,8–12,14,19]. The Handbook of Computational Social Choice provides an excellent survey of this broad field, which we cannot overview in its entirety [15]. Our work considers the model initially posed in [1,2], and continued in [6,17,18], which we overview in more detail below.

[1,2] were the first to consider Tournament rules that are both Condorcet-consistent and 2-SNM, and proved that no such rules exist. They further considered tournament rules that are 2-SNM and approximately Condorcet-consistent. [17] first proposed to consider tournament rules that are instead Condorcet-consistent and approximately 2-SNM. Their work establishes that Randomized Single Elimination Bracket is 2-SNM-1/3, and that this is tight.[2] [18] establish that Randomized King of the Hill (RKotH) is 2-SNM-1/3,[3] and [6] establish that Randomized Death Match is 2-SNM-1/3. [18] show further that RKotH satisfies a stronger fairness notion called Cover-consistence, and [6] extends their analysis to probabilistic tournaments. In summary, the state of affairs for $k = 2$ is quite established: multiple 2-SNM-1/3 tournament rules are known, and multiple different extensions beyond the initial model of [17] are known.

For $k > 2$, however, significantly less is known. [17] gives a simple example establishing that no rule is k-SNM-$\frac{k-1-\varepsilon}{2k-1}$ for any $\varepsilon > 0$, but no rules are known to match this bound for any $k > 2$. Indeed, [18] shows that this bound is not tight, and proves a stronger lower bound for $k \to \infty$. For example, a corollary of their main result is that no 939-SNM-1/2 tournament rule exists. They also design a non-monotone tournament rule that is k-SNM-2/3 for all k. Other than these results, there is no prior work for manipulating sets of size $k > 2$.

[2] Randomized Single Elimination Bracket iteratively places the teams, randomly, into a single-elimination bracket, and then 'plays' all matches that would occur in this bracket to determine a winner.

[3] Randomized King of the Hill iteratively picks a 'prince', and eliminates all teams beaten by the prince, until only one team remains.

In comparison, our work is the first to give a tight analysis of any Condorcet-consistent tournament rule for $k > 2$, and is the first proof that any monotone, Condorcet-consistent tournament rule is k-SNM-α for any $k > 2, \alpha < 1$.

Regarding our study of Sybil attacks, similar clone manipulations have been considered prior in Social Choice Theory under the name of *composition-consistency*. [13] introduces the notion of a *decomposition* of the teams in a tournament into components, where all the teams in a component are clones of each other with respect to the teams not in the component. [13] defines a deterministic tournament rule to be *composition-consistent* if it chooses the best teams from the best components[4]. In particular, composition-consistency implies that a losing team cannot win by introducing clones of itself or any other team. [13] shows that the tournament rules Banks, Uncovered Set, TEQ, and Minimal Covering Set are *composition-consistent*, while Top Cycle, the Slater, and the Copeland are not. Both computational and axiomatic aspects of *composition-consistency* have been explored ever since. [7] studies the structural properties of clone sets and their computational aspects in the context of voting preferences. In the context of probabilistic social choice, [4] gives probabilistic extensions of the axioms *composition-consistency* and *population-consistency* and uniquely characterize the probabilistic social choice rules, which satisfy both. In the context of scoring rules, [16] studies the incompatibility of *composition-consistency* and *reinforcement* (stronger than *population-consistency*) and decomposes composition-consistency into four weaker axioms. In this work, we consider Sybil attacks on Randomized Death Match. Our study of Sybil attacks differs from prior work on the relevant notion of *composition-consistency* in the following ways: (i) We focus on a randomized tournament rule (RDM), (ii) We study settings where the manipulator creates clones of themselves (i.e. not of other teams), (iii) We explore the asymptotic behavior of such manipulations (Definition 10, Theorem 8).

1.4 Roadmap

Section 2 follows with definitions and preliminaries, and formally defines Randomized Death Match (RDM). Section 3 introduces some basic properties and examples for the RDM rule as well as a recap of previous work for two manipulators. Section 4 consists of a proof that the manipulability of 3 teams in RDM is at most $\frac{31}{60}$ and that this bound is tight. Section 5 consists of our main results regarding Sybil attacks on a tournament. Section 6 concludes.

2 Preliminaries

In this section we introduce notation that we will use throughout the paper consistent with prior work in [6,17,18].

[4] For a full rigorous mathematical definition see Definition 10, [13].

Definition 1 (Tournament). *A (round robin) tournament T on n teams is a complete, directed graph on n vertices whose edges denote the outcome of a match between two teams. Team i beats team j if the edge between them points from i to j.*

Definition 2 (Tournament Rule). *A tournament rule r is a function that maps tournaments T to a distribution over teams, where $r_i(T) := \Pr(r(T) = i)$ denotes the probability that team i is declared the winner of tournament T under rule r. Let S be a set of teams. We use the shorthand $r_S(T) := \sum_{i \in S} r_i(T)$ to denote the probability that a team in S is declared the winner of tournament T under rule r.*

Definition 3 (S-adjacent). *Let S be a set of teams. Two tournaments T, T' are S-adjacent if for all i, j such that $\{i, j\} \not\subseteq S$, i beats j in T if and only if i beats j in T'.*

In other words, two tournaments T, T' are S-adjacent if the teams from S can manipulate the outcomes of the matches between them in order to obtain a new tournament T'.

Definition 4 (Condorcet-Consistent). *Team i is a Condorcet winner of a tournament T if i beats every other team (under T). A tournament rule r is Condorcet-consistent if for every tournament T with a Condorcet winner i, $r_i(T) = 1$ (whenever T has a Condorcet winner, that team wins with probability 1).*

Definition 5 (Manipulating a Tournament). *For a set S of teams, a tournament T and a tournament rule r, we define $\alpha_S^r(T)$ be the maximum winning probability that S can possibly gain by manipulating T to an S-adjacent T'. That is:*

$$\alpha_S^r(T) = \max_{T' : T' \text{ is } S\text{-adjacent to } T} \{r_S(T') - r_S(T)\}$$

For a tournament rule r, define $\alpha_{k,n}^r = \sup_{T,S:|S|=k,|T|=n}\{\alpha_S^r(T)\}$. Finally, define

$$\alpha_k^r = \sup_{n \in \mathbb{N}} \alpha_{k,n}^r = \sup_{T,S:|S|=k} \{\alpha_S^r(T)\}$$

If $\alpha_k^r \leq \alpha$, we say that r is k-Strongly-Non-Manipulable at probability α or k-SNM-α.

Intuitively, $\alpha_{k,n}^r$ is the maximum increase in collective winning probability that a group of k teams can achieve by manipulating the matches between them over tournaments with n teams. And α_k^r is the maximum increase in winning probability that a group of k teams can achieve by manipulating the matches between them over all tournaments.

Two other naturally desirable properties of a tournament rule are monotonicity and anonymity.

Definition 6 (Monotone). *A tournament rule r is monotone if T, T' are $\{u, v\}$-adjacent and u beats v in T, then $r_u(T) \geq r_u(T')$*

Definition 7 (Anonymous). *A tournament rule r is anonymous if for every tournament T, and every permutation σ, and all i, $r_{\sigma(i)}(\sigma(T)) = r_i(T)$*

Below we define the tournament rule that is the focus of this work.

Definition 8 (Randomized Death Match). *Given a tournament T on n teams the Randomized Death Match Rule (RDM) picks two uniformly random teams (without replacement) and plays their match. Then, eliminates the loser and recurses on the remaining teams for a total of $n-1$ rounds until a single team remains, who is declared the winner.*

Below we define the notions of *Sybil Attack* on a tournament T, and the property of *Asymptotically Strongly Sybil-Proof* (ASSP) for a tournament rule r, both of which will be relevant in our discussion in Sect. 5.

Definition 9 (Sybil Attack). *Given a tournament T, a team $u_1 \in T$ and an integer m, define $Syb(T, u_1, m)$ to be the set of tournaments T' satisfying the following properties:*

1. *The set of teams in T' consists of $u_2 \dots, u_m$ and all teams in T*
2. *If a, b are teams in T, then a beats b in T' if and only if a beats b in T.*
3. *If $a \neq u_1$ is a team in T and $i \in [m]$, then u_i beats a in T' if and only if u_1 beats a in T*
4. *The match between u_i and u_j can be arbitrary for each $i \neq j$*

Intuitively, $Syb(T, u_1, m)$ is the set of all Sybil attacks of u_1 at T with m Sybils. Each Sybil attack is a tournament $T' \in Syb(T, u_1, m)$ obtained by starting from T and creating m Sybils of u_1 (while counting u_1 as a Sybil of itself). Each Sybil beats the same set of teams from $T \backslash u_1$ and the matches between the Sybils $u_1, \dots, u_m$ can be arbitrary. Every possible realization of the matches between the Sybils gives rise to new tournament $T' \in Syb(T, u_1, m)$ (implying $Syb(T, u_1, m)$ contains $2^{\binom{m}{2}}$ tournaments).

Definition 10 (Asymptotically Strongly Sybil-Proof). *A tournament rule r is Asymptotically Strongly Sybil-Proof (ASSP) if for any tournament T, team $u_1 \in T$ which is not a Condorcet winner,*

$$\lim_{m \to \infty} \max_{T' \in Syb(T, u_1, m)} r_{u_1, \dots, u_m}(T') = 0$$

Informally speaking, Definition 10 claims that r is ASSP if the probability that a Sybil wins in the most profitable Sybil attack on T with m Sybils, goes to zero as m goes to ∞.

3 Basic Properties of RDM and Examples

In this section we consider a few basic properties of RDM and several examples on small tournaments. We will refer to those examples in our analysis later.

Throughout the paper we will denote RDM by r and it will be the only tournament rule we consider. We next state the first-step analysis observation that will be central to our analysis throughout the paper. For the remainder of the section let for a match e denote by $T|_e$ the tournament obtained from T by eliminating the loser in e. Let $S|_e = S\backslash x$, where x is the loser in e. Let d_x denote the number of teams x loses to and $T\backslash x$ the tournament obtained after removing team x from T.

Observation 3 (First-step analysis). *Let S be a subset of teams in a tournament T. Then*

$$r_S(T) = \frac{1}{\binom{n}{2}} \sum_e r_{S|_e}(T|_e) = \frac{1}{\binom{n}{2}} \sum_x d_x r_{S\backslash x}(T\backslash x)$$

(if $S = \{v\}$, then we define $r_{S\backslash v}(T\backslash v) = 0$, and if $x \notin S$, then $S\backslash x = S$)

Proof. The first equality follows from the fact that after we play e we are left with the tournament $T|_e$ and we sum over all possible e in the first round. To prove the second equality, notice that for any x the term $r_{S\backslash x}(T\backslash x)$ appears exactly d_x times in $\sum_e r_{S|_e}(T|_e)$ because x loses exactly d_x matches.

This first-step analysis observation can be used to show that adding teams that lose to every other team does not change the probability distribution of the winner.

Lemma 1. *Let T be a tournament and $u \in T$ loses to everyone. Then for all $v \neq u$, we have $r_v(T) = r_v(T\backslash u)$.*

Proof. See full version [5] for a proof.

As a natural consequence of Lemma 1 we show that the most manipulable tournament on $n+1$ teams is at least as manipulable as the most manipulable tournament on n teams.

Lemma 2. $\alpha^r_{k,n} \leq \alpha^r_{k,n+1}$

Proof. See Appendix A.1 of [5] for a proof.

We now show another natural property of RDM, which is a generalization of Condorcet-consistent (Definition 4), namely that if a group of teams S wins all its matches against the rest of teams, then a team from S will always win.

Lemma 3. *Let T be a tournament and $S \subseteq T$ a group of teams such that every team in S beats every team in $T\backslash S$. Then, $r_S(T) = 1$.*

Proof. See Appendix A.1 of [5] for a proof.

As a result of Lemma 3 RDM is Condorcet-Consistent. As expected, RDM is also monotone (See Definition 6).

Lemma 4. *RDM is monotone.*

Proof. See Appendix A.1 of [5] for a proof.

Lemma 1 tells us that adding a team which loses to all other teams does not change the probability distribution of the other teams winning. Lemmas 1, 2, 3, 4 will be useful in our later analysis in Sects. 4 and 5. Now we consider a few examples of tournaments and illustrate the use of first-step analysis (Observation 3) to compute the probability distribution of the winner in them.

1. Let $T = \{a, b, c\}$, where a beats b, b beats c and c beats a. By symmetry of RDM, we have $r_a(T) = r_b(T) = r_c(T) = \frac{1}{3}$.
2. Let $T = \{a, b, c\}$ where a beats b and c. Then clearly, $r_a(T) = 1$ and $r_b(T) = r_c(T) = 0$.
3. By Lemma 1, it follows that the only tournament on 4 teams whose probability distribution cannot be reduced to a distribution on 3 teams is the following one $T = \{a_1, a_2, a_3, a_4\}$, where a_i beats a_{i+1} for $i = 1, 2, 3$, a_4 beats a_1, a_1 beats a_3 and a_2 beats a_4. By using what we computed in (1) and (2) combined with Lemma 1 we get by first step analysis

$$r_{a_1}(T) = \frac{1}{6}(r_{a_1}(T \backslash a_2) + 2r_{a_1}(T \backslash a_3) + 2r_{a_1}(T \backslash a_4)) = \frac{1}{6}(\frac{1}{3} + \frac{2}{3} + 2) = \frac{1}{2}$$
$$r_{a_2}(T) = \frac{1}{6}(r_{a_2}(T \backslash a_1) + 2r_{a_2}(T \backslash a_3) + 2r_{a_2}(T \backslash a_4)) = \frac{1}{6}(1 + \frac{2}{3}) = \frac{5}{18}$$
$$r_{a_3}(T) = \frac{1}{6}(r_{a_3}(T \backslash a_1) + r_{a_3}(T \backslash a_2) + 2r_{a_3}(T \backslash a_4)) = \frac{1}{6}(\frac{1}{3}) = \frac{1}{18}$$
$$r_{a_4}(T) = \frac{1}{6}(r_{a_4}(T \backslash a_1) + r_{a_4}(T \backslash a_2) + 2r_{a_4}(T \backslash a_3)) = \frac{1}{6}(\frac{1}{3} + \frac{2}{3}) = \frac{1}{6}$$

The above examples are really important in our analysis because: a) we will use them in later for our lower bound example in Sect. 4.1, and b) they are a short illustration of first-step analysis.

In the following subsection, we review prior results for 2-team manipulations in RDM, which will also be useful for our treatment of the main result in Sect. 4.

3.1 Recap: Tight Bounds on 2-Team Manipulations in RDM

[6] (Theorem 5.2) proves that RDM is 2-SNM-$\frac{1}{3}$ and that this bound is tight, namely $\alpha_2^{RDM} = \frac{1}{3}$. We will rely on this result in Sect. 4.

Theorem 4 *(Theorem 5.2 in [6]).* $\alpha_2^{RDM} = \frac{1}{3}$

[17] (Theorem 3.1) proves that the bound of $\frac{1}{3}$ is the best one can hope to achieve for a Condorcet-consistent rule.

Theorem 5 *(Theorem 3.1 in [17]). There is no Condorcet-consistent tournament rule on n players (for $n \geq 3$) that is 2-SNM-α for $\alpha < \frac{1}{3}$*

We prove the following useful corollary, which will be useful in Sect. 4.

Corollary 1. *Let T be a tournament and $u, v \in T$ two teams such that there is at most one match in which a team in $\{u, v\}$ loses to a team in $T \backslash \{u, v\}$. Then*

$$r_{u,v}(T) \geq \frac{2}{3}$$

Proof. See full version [5] for proof.

4 Main Result: $\alpha_3^{RDM} = 31/60$

The goal of this section is to prove that no 3 teams can improve their probability of winning by more than $\frac{31}{60}$ and that this bound is tight. We prove the following theorem

Theorem 6. $\alpha_3^{RDM} = \frac{31}{60}$

Our proof consists of two parts:

- Lower bound: $\alpha_3^{RDM} \geq \frac{31}{60}$, for which we provide a tournament T and a set S of size 3, which can manipulate to increase their probability by $\frac{31}{60}$
- Upper bound: $\alpha_3^{RDM} \leq \frac{31}{60}$, for which we provide a proof that for any tournament T no set S of size 3 can increase their probability of winning by more than $\frac{31}{60}$, i.e. RDM is 3-SNM-$\frac{31}{60}$

4.1 Lower Bound

Let r denote RDM. Denote by B_x the set of teams which team x beats. Consider the following tournament $T = \{u, v, w, a, b\}$:

$$B_a = \{u, v, b\}, B_b = \{u, v\}, B_u = \{v, w\}, B_v = \{w\}, B_w = \{a, b\}$$

Let $S = \{u, v, w\}$. By first-step analysis (Observation 3) and by using our knowledge in Sect. 3 for tournaments on 4 teams we can show that $r_{u,v,w}(T) = \frac{29}{60}$ (see full version [5] for full analysis) Let u and v throw their matches with w. i.e. T' is S-adjacent to T, where in T', w beats u and v and all other matches have the same outcomes as in T. Then, since w is a Condorcet winner, $r_{u,v,w}(T') = r_w(T') = 1$. Therefore,

$$\alpha_3^{RDM} \geq r_{u,v,w}(T') - r_{u,v,w}(T) = 1 - \frac{29}{60} = \frac{31}{60}$$

Thus, $\alpha_3^{RDM} \geq \frac{31}{60}$ as desired.

4.2 Upper Bound

Suppose we have a tournament T on $n \geq 3$ vertices and $S = \{u, v, w\}$ is a set of 3 (distinct) teams, where S will be the set of manipulator teams. Let I be the set of matches in which a team from S loses to a team from $T \backslash S$. Our proof for $\alpha_k^{RDM} \leq \frac{31}{60}$ will use the following strategy

- In the following section we introduce the first-step analysis framework by considering possible cases for the first played match. In each of these cases the loser of the match is eliminated and we are left with a tournament with one less team. We pair each match in T with its corresponding match in T' and we bound the gains of manipulation in each of the following cases separately (these correspond to each of the terms A, B, and C respectively in the analysis in the next section).
  - The first match is between two teams in S (there are 3 such matches).
  - The first match is between a team in S and a team in $T \backslash S$ and the team from S loses in the match (there are $|I|$ such matches).
  - The first match is any other match not covered by the above two cases
- We prove that if $|I| \leq 4$, then $\alpha_S^{RDM}(T) \leq \frac{31}{60}$ (i.e. the set of manipulators cannot gain more than $\frac{31}{60}$ by manipulating).
- We prove that if T is the most manipulable tournament on n vertices (i.e. $\alpha_S^{RDM}(T) = \alpha_{3,n}^{RDM}$), then $\alpha_S^{RDM}(T) \leq \frac{|I|+7}{3(|I|+3)}$
- We combine the above facts to finish the proof of Theorem 6

We first introduce some notation that we will use throughout this section. Suppose that T' and T are S-adjacent. Recall from Sect. 3 that for a match $e = (i, j)$, $T|_e$ is the tournament obtained after eliminating the loser in e. Also d_x is the number of teams that a team x loses to in T. For $x \in S$, let ℓ_x denote the number of matches x loses against a team in S when considered in T and let ℓ'_x denote the number of matches that x loses against a team in S when considered in T'. Let d_x^* denote the number of teams in $T \backslash S$ that x loses to. Notice that since T and T' are S-adjacent, $x \in S$ loses to exactly d_x^* teams in $T' \backslash S$ when considered in T'. Let $G = I \cup \{uv, vw, uw\}$ be the set of matches in which a team from S loses.

The First Step Analysis Framework. Notice that in the first round of RDM, a uniformly random match e from the $\binom{n}{2}$ matches is chosen. If $e \in G$ then we are left with $T \backslash x$ where x loses in e for some $x \in S$. If $e \notin G$, we are left with $T|_e$ and all teams in S are still in the tournament. For $x \in S$, there are ℓ_x matches in which they lose to a team from S and d_x^* matches in which they lose to a team from $T \backslash S$. We consider each of these cases and use first-step analysis (Observation 3) for both T and T' to obtain (see full version [5] for details)

$$r_{u,v,w}(T') - r_{u,v,w}(T) = \frac{1}{\binom{n}{2}}(A + B + C) \tag{1}$$

where

$$A = \sum_{x \in \{u,v,w\}} \ell'_x r_{\{u,v,w\} \backslash x}(T' \backslash x) - \ell_x r_{\{u,v,w\} \backslash x}(T \backslash x)$$

$$B = \sum_{x \in S} d_x^* (r_{\{u,v,w\} \backslash x}(T' \backslash x) - r_{\{u,v,w\} \backslash x}(T \backslash x))$$

$$C = \sum_{e \notin G} r_{u,v,w}(T'|_e) - r_{u,v,w}(T|_e)$$

Upper Bounds on $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$. We now prove some bounds on the terms A, B and C (defined in Sect. 4.2) which will be useful later. Recall that I denotes the set of matches in which a team from S loses from a team from $T \backslash S$. We begin with bounding A in the following lemma

Lemma 5. *For all S-adjacent T and T', we have $A \leq \frac{7}{3}$. Moreover, if $|I| \leq 1$, then $A \leq 1$.*

Proof. See Appendix A.2 in [5] for a proof.

Next, we show the following bound on the term B.

Lemma 6. *For all S-adjacent T, T' we have*

$$B \leq \frac{d_u^* + d_v^* + d_w^*}{3} = \frac{|I|}{3}$$

Moreover, if $|I| \leq 1$, then $B = 0$

Proof. See Appendix A.2 in [5] for a proof.

We introduce some more notation. For $n \in \mathbb{N}$, define $M_n(a_1, a_2, a_3)$ as the maximum winning probability gain that three teams $\{u, v, w\}$ can achieve by manipulation in a tournament T of size n, in which there are exactly a_i teams in $T \backslash S$ each of which beats exactly i teams of S. Formally,

$$\begin{aligned} M_n(a_1, a_2, a_3) = \max \Big\{ & r_S(T') - r_S(T) | T, T' \text{ are } S\text{-adjacent}, |T| = n, |S| = 3, \\ & a_i \text{ teams in } T \backslash S \text{ beat exactly } i \text{ teams in } S \Big\} \end{aligned}$$

Additionally, let L_i be the set of teams in $T \backslash S$ each of which beats exactly i teams in S. Let Q be the set of matches in which two teams from L_i play against each other or in which a team from L_i loses to a team from S for $i = 1, 2, 3$. Notice that $|Q| = 2a_1 + a_2 + \binom{a_1}{2} + \binom{a_2}{2} + \binom{a_3}{2}$ if there are a_i teams in $S \backslash T$ each of which beat i teams from S.

With the new notation, we are now ready to prove a bound on the term C. Recall that

$$C = \sum_{e \notin G} r_{u,v,w}(T'|_e) - r_{u,v,w}(T|_e)$$

where $G = I \cup \{uv, vw, uw\}$ is the set of matches in which a team from S loses. Then we have the following bound on C.

Lemma 7. *For all S-adjacent T and T' we have that C is at most*

$$\begin{aligned} &(2a_1 + \binom{a_1}{2}) M_{n-1}(a_1 - 1, a_2, a_3) + (a_2 + \binom{a_2}{2}) M_{n-1}(a_1, a_2 - 1, a_3) \\ &+ \binom{a_3}{2} M_n(a_1, a_2, a_3 - 1) + \sum_{e \notin G \cup Q} r_{u,v,w}(T'|_e) - r_{u,v,w}(T|_e) \end{aligned}$$

Proof. See Appendix A.2 in [5] for a proof.

The Case $|I| \leq 4$. We summarize our claim when $|I| \leq 4$ in the following lemma

Lemma 8. *Let T be a tournament, and S a set of 3 teams. Suppose that there are at most 4 matches in which a team in S loses to a team in $T \backslash S$ (i.e. $|I| \leq 4$). Then $\alpha_S^{RDM}(T) \leq \frac{31}{60}$*

Proof. We will show that $M_n(a_1, a_2, a_3) \leq f(a_1, a_2, a_3)$ by induction on $n \in \mathbb{N}$ for the values of (a_1, a_2, a_3) and $f(a_1, a_2, a_3)$ given in Table 1 below. Notice that when there are at most 4 matches between a team in S and a team in $T \backslash S$, in which the teams from S loses, then we fall into one of the cases shown in the table for (a_1, a_2, a_3).

Table 1. Upper bounds on $M_n(a_1, a_2, a_3)$

(a_1, a_2, a_3)	$f(a_1, a_2, a_3)$
(0, 0, 0)	0
(1, 0, 0)	$\frac{1}{6}$
(2, 0, 0)	$\frac{23}{60}$
(3, 0, 0)	$\frac{407}{900}$
(4, 0, 0)	$\frac{4499}{9450}$
(0, 1, 0)	$\frac{1}{2}$
(0, 2, 0)	$\frac{31}{60}$
(1, 1, 0)	$\frac{1}{2}$
(2, 1, 0)	$\frac{131}{260}$
(0, 0, 1)	0
(1, 0, 1)	$\frac{11}{27}$

1. Base case. Our base case is when $n = 3$. If we are in the case of 3 teams then S wins with probability 1, so the maximum gain S can achieve by manipulation is clearly 0, which satisfies all of the bounds in the table.

2. Induction step. Assume that $M_k(a_1, a_2, a_3) \leq f(a_1, a_2, a_3)$ hold for all $k < n$ and a_1, a_2, a_3 as in Table 1. We will prove the statement for $k = n$. By using the upper bounds on A, B, C in Lemmas 5–7 and the inductive hypothesis we can show that (see [5] for full analysis)

$$M_n(a_1,a_2,a_3) \leq \frac{1}{\binom{n}{2}}\Bigg[A' + B' + (2a_1 + \binom{a_1}{2})f(a_1-1,a_2,a_3)$$
$$+ (a_2 + \binom{a_2}{2})f(a_1,a_2-1,a_3) + \binom{a_3}{2}f(a_1,a_2,a_3-1)$$
$$+ (\binom{n}{2} - 3(1+a_1+a_2+a_3) - \binom{a_1}{2} - \binom{a_2}{2} - \binom{a_3}{2})f(a_1,a_2,a_3)\Bigg] \quad (\,\Delta\,)$$

where $A' = 1$ and $B' = 0$ if $|I| \leq 1$ and $A' = \frac{7}{3}$ and $B' = \frac{|I|}{3}$ if $|I| \geq 2$. Next, we apply the formula (Δ) to each of the cases in Table 1, to obtain $M_n(a_1,a_2,a_3) \leq f(a_1,a_2,a_3)$. We defer the reader to [5] for full details.

This finishes the induction and the proof for the bounds in Table 1. Note that $f(a_1,a_2,a_3) \leq \frac{31}{60}$ for all a_1,a_2,a_3 in Table 1 and this bounds is achieved when $(a_1,a_2,a_3) = (0,2,0)$ i.e. there are 2 teams that beat exactly two of S as is the case in the optimal example in Sect. 4.1. Thus,we get that if there are at most 4 matches that a team from S loses from a team in $T \backslash S$, then $\alpha_S^{RDM}(T) \leq \frac{31}{60}$. This finishes the proof of the lemma.

General Upper Bound for the Most Manipulable Tournament

Lemma 9. *Suppose that* $\alpha_S^{RDM}(T) = \alpha_{3,n}^{RDM}$. *Let* I *be the set of matches a team of* S *loses to a team from* $T \backslash S$. *Then*

$$\alpha_{3,n}^{RDM} = \alpha_S^{RDM}(T) \leq \frac{|I|+7}{3(|I|+3)}$$

Proof. Let T and T' be S-adjacent tournaments on n vertices such that $S = \{u,v,w\}$ and
$$\alpha_{3,n}^{RDM} = \alpha_S^{RDM}(T) = r_S(T') - r_s(T)$$
I.e. T is the "worst" example on n vertices. By (1) we have

$$\alpha_{3,n}^{RDM} = \frac{1}{\binom{n}{2}}(A+B+C)$$

where A, B and C were defined in Sect. 4.2. By Lemma 5 we have

$$A \leq \frac{7}{3}$$

and by Lemma 6
$$B \leq \frac{d_u^* + d_v^* + d_w^*}{3} = \frac{|I|}{3}$$

Let $e \notin G$. Notice that both $T'|_e$ and $T|_e$ are tournaments on $n-1$ vertices and by definition of G, u, v, w are not eliminated in both $T'|_e$ and $T|_e$. Moreover, $T'|_e$ and $T|_e$ are S-adjacent. Therefore, for every $e \notin G$, we have by Lemma 2

$$r_{u,v,w}(T'|_e) - r_{u,v,w}(T|_e) \leq \alpha_{3,n-1}^{RDM} \leq \alpha_{3,n}^{RDM}$$

By using the above on each term in C and the fact that $|G| = 3 + |I|$, we get that

$$C \leq (\binom{n}{2} - (3 + |I|))\alpha_{3,n}^{RDM}$$

By using the above 3 bounds we get

$$\begin{aligned} \alpha_{3,n}^{RDM} &\leq \frac{1}{\binom{n}{2}}(\frac{7}{3} + \frac{|I|}{3} + (\binom{n}{2} - (3 + |I|))\alpha_{3,n}^{RDM}) \\ \iff (|I| + 3)\alpha_{3,n}^{RDM} &\leq \frac{|I| + 7}{3} \\ \iff \alpha_{3,n}^{RDM} = \alpha_S^{RDM}(T) &\leq \frac{|I| + 7}{3(|I| + 3)} \end{aligned}$$

as desired.

Proof of Theorem 6. Suppose that T is the most manipulable tournament on n vertices i.e. it satisfies $\alpha_S^{RDM}(T) = \alpha_{3,n}^{RDM}$. If $|I| \leq 4$, then by Lemma 8, we have that

$$\alpha_{3,n}^{RDM} = \alpha_S^{RDM}(T) \leq \frac{31}{60}$$

If $|I| \geq 5$, then by Lemma 9

$$\alpha_{3,n}^{RDM} = \alpha_S^{RDM}(T) \leq \frac{|I| + 7}{3(|I| + 3)} \leq \frac{5 + 7}{3(5 + 3)} = \frac{1}{2}$$

where above we used that $\frac{x+7}{3(x+3)}$ is decreasing for $x \geq 5$. Combining the above bounds, we obtain $\alpha_{3,n}^{RDM} \leq \frac{31}{60}$ for all $n \in \mathbb{N}$. Therefore,

$$\alpha_k^{RDM} = \max_{n \in \mathbb{N}} \alpha_{k,n}^{RDM} \leq \frac{31}{60}$$

which proves the upper bound and finishes the proof of Theorem 6.

5 Sybil Attacks on Tournaments

5.1 Main Results on Sybil Attacks on Tournaments

Recall our motivation from the Introduction. Imagine that a tournament rule is used as a proxy for a voting rule to select a proposal. The proposals are compared head-to-head, and this constitutes the pairwise matches in the resulting tournament. A proposer can try to manipulate the protocol with a Sybil attack and submit many nearly identical proposals with nearly equal strength relative to the other proposals. The proposer can choose to manipulate the outcomes of the head-to-head comparisons between two of his proposals in a way which maximizes the probability that a proposal of his is selected. In the tournament T his proposal corresponds to a team u_1, and the tournament T' resulting from the Sybil attack is a member of $Syb(T, u_1, m)$ (Recall Definition 9). The questions that we want to answer in this section are: (1) Can the Sybils manipulate their matches to successfully increase their collective probability of winning? and (2) Is it beneficial for the proposer to create as many Sybils as possible?

The first question we are interested in is whether any group of Sybils can manipulate successfully to increase their probability of winning. It turns out that the answer is No. We first prove that the probability that a team that is not a Sybil wins does not depend on the matches between the Sybils.

Lemma 10. *There exists a function q that takes in as input integer m, tournament T, team $u_1 \in T$, and team $v \in T\backslash u_1$ with the following property. For all $T' \in Syb(T, u_1, m)$, we have*

$$r_v(T') = q(m, T, u_1, v)$$

where the dependence on u_1 is encoded as the outcomes of its matches with the rest of T.

Proof. See Appendix A.3 in [5] for a full proof.

Note that by Lemma 10 $r_v(T') = q(m, T, u_1, v)$ does not depend on which tournament $T' \in Syb(T, u_1, m)$ is chosen. Now, we prove our first promised result. Namely, that no number of Sybils in a Sybil attack can manipulate the matches between them to increase their probability of winning.

Theorem 7. *Let T be a tournament, $u_1 \in T$ a team, and m and integer. Let $T_1' \in Syb(T, u_1, m)$. Let $S = \{u_1, \ldots, u_m\}$. Then*

$$\alpha_S^{RDM}(T_1') = 0$$

Proof. Let T_1' and T_2' be S-adjacent. By Definition 9, $T_2' \in Syb(T, u_1, m)$. Therefore by Lemma 10, $r_v(T_1') = r_v(T_2') = q(m, T, u_1, v)$ for all $v \in T\backslash u_1$. Using this we obtain

$$r_S(T_1') = 1 - \sum_{v \in T\backslash u_1} r_v(T_1') = 1 - \sum_{v \in T\backslash u_1} r_v(T_2') = r_S(T_2')$$

Therefore, $r_S(T_1') = r_S(T_2')$ for all S-adjacent T_1', T_2', which implies the desired result.

We are now ready to prove our second main result. Namely, that RDM is Asymptotically Strongly Sybil-Proof (Definition 10). Before we present the result (Theorem 8) we will try to convey some intuition for why RDM should be ASSP.

Let u_1 be a team to create Sybils of (u_1 is not Condorcet Winner). Let A be the set of teams which u_1 beats and B the set of teams which u_1 loses to in T. Observe that the only way a Sybil can win is when all the teams from B are eliminated before all the Sybils. The teams from B can only be eliminated by teams from A. However, as m increases there are more Sybils and, thus, the teams from A are intuitively more likely to all lose the tournament before the teams from B. When there are no teams from A left and at least one team from B left, no Sybil can win. In fact, this intuition implies something stronger than RDM being ASSP: the collective winning probability of the Sybils and the teams from A converges to 0 as m converges to ∞ (or, equivalently, the probability that a team from B wins goes to 1). This intuition indirectly lies behind the technical details of the proof of Theorem 8. Let $p(m, T, u_1)$ be the collective probability that any Sybil or a team in A wins. It is not hard to see that by Lemma 10 it doesn't depend on the matches between the Sybils (see [5] for full details). Then we have the following theorem.

Theorem 8. *Randomized Death Match is Asymptotically Strongly Sybil-Proof. In fact a stronger statement holds, namely if $u_1 \in T$ is not a Condorcet winner, then*

$$\lim_{m\to\infty} p(m, T, u_1) = 0$$

Proof. See Appendix A.3 in [5] for a full proof.

5.2 On a Counterexample to an Intuitive Claim

We will use Theorem 8 to prove that RDM does not satisfy a stronger version of the monotonicity property in Definition 6. First, we give a generalization of the definition for monotonicity given in Sect. 3

Definition 11 (Strongly monotone). *Let r be a tournament rule. Let T and let $C \cup D$ be any partition of the teams in T into two disjoint sets. A tournament rule r is strongly monotone for every $(u, v) \in C \times D$, if T' is $\{u, v\}$-adjacent to T such that u beats v in T' we have $r_C(T') \geq r_C(T)$*

Intuitively, r is Strongly monotone if whenever flipping a match between a team from C and a team from D in favor of the team from C makes C better off. Notice that if $|C| = 1$ this is the usual definition of monotonicity (Definition 6), which is satisfied by RDM by Lemma 4. However, RDM is not strongly monotone, even though strong monotonicity may seem like an intuitive property to have.

Theorem 9. *RDM is not strongly monotone*

Proof. See [5] for a proof.

6 Conclusion and Future Work

We use a novel first-step analysis to nail down the minimal α such that RDM is 3-SNM-α. Specifically, our main result shows that $\alpha_3^{RDM} = \frac{31}{60}$. Recall that this is the first tight analysis of any Condorcet-consistent tournament rule for any $k > 2$, and also the first monotone, Condorcet-consistent tournament rule that is k-SNM-α for any $k > 2, \alpha < 1$. We also initiate the study of manipulability via Sybil attacks, and prove that RDM is Asymptotically Strongly Sybil-Proof.

Our technical approach opens up the possibility of analyzing the manipulability of RDM (or other tournament rules) whose worst-case examples are complicated-but-tractable. For example, it is unlikely that the elegant coupling arguments that work for $k = 2$ will result in a tight bound of 31/60, but our approach is able to drastically reduce the search space for a worst-case example, and a tractable case analysis confirms that a specific 5-team tournament is tight. Our approach can similarly be used to analyze the manipulability of RDM for $k > 3$, or other tournament rules. However, there are still significant technical barriers for future work to overcome in order to keep analysis tractable for large k, or for tournament rules with a more complicated recursive step. Still, our techniques provide a clear approach to such analyses that was previously non-existent.

References

1. Altman, A., Kleinberg, R.: Nonmanipulable randomized tournament selections. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 24 (2010)
2. Altman, A., Procaccia, A.D., Tennenholtz, M.: Nonmanipulable selections from a tournament. In: Computational Foundations of Social Choice. Dagstuhl Seminar Proceedings, 07 March–12 March 2010, vol. 10101. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany (2010). http://drops.dagstuhl.de/opus/volltexte/2010/2560/
3. Banks, J.S.: Sophisticated voting outcomes and agenda control. Soc. Choice Welfare **1**(4), 295–306 (1985). https://doi.org/10.1007/BF00649265. http://www.jstor.org/stable/41105790
4. Brandl, F., Brandt, F., Seedig, H.G.: Consistent probabilistic social choice. Econometrica **84**(5), 1839–1880 (2016)
5. Dinev, A., Weinberg, S.M.: Tight bounds on 3-team manipulations in randomized death match. In: Hansen, K.A., et al. (eds.) WINE 2022. LNCS, vol. 13778, pp. 273–291. Springer, Cham (2022)
6. Ding, K., Weinberg, S.M.: Approximately strategyproof tournament rules in the probabilistic setting. In: 12th Innovations in Theoretical Computer Science Conference, ITCS (2021)
7. Elkind, E., Faliszewski, P., Slinko, A.: Clone structures in voters' preferences. In: Proceedings of the 13th ACM Conference on Electronic Commerce, pp. 496–513 (2012)
8. Fishburn, P.C.: Condorcet social choice functions. SIAM J. Appl. Math. **33**(3), 469–489 (1977). http://www.jstor.org/stable/2100704
9. Fisher, D.C., Ryan, J.: Optimal strategies for a generalized "scissors, paper, and stone" game. Am. Math. Monthly **99**(10), 935–942 (1992). http://www.jstor.org/stable/2324486

10. Laffond, G., Laslier, J.F., Le Breton, M.: The bipartisan set of a tournament game. Games Econ. Behav. **5**(1), 182–201 (1993). https://EconPapers.repec.org/RePEc:eee:gamebe:v:5:y:1993:i:1:p:182-201
11. Kim, M.P., Suksompong, W., Williams, V.V.: Who can win a single-elimination tournament? SIAM J. Discrete Math. **31**(3), 1751–1764 (2017)
12. Kim, M.P., Williams, V.V.: Fixing tournaments for kings, chokers, and more. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
13. Laffond, G., Lainé, J., Laslier, J.F.: Composition-consistent tournament solutions and social choice functions. Soc. Choice Welfare **13**(1), 75–93 (1996). https://doi.org/10.1007/BF00179100
14. Miller, N.R.: A new solution set for tournaments and majority voting: further graph-theoretical approaches to the theory of voting. Am. J. Polit. Sci. **24**(1), 68–96 (1980). http://www.jstor.org/stable/2110925
15. Moulin, H.: Handbook of Computational Social Choice. Cambridge University Press, Cambridge (2016). https://doi.org/10.1017/CBO9781107446984
16. Öztürk, Z.E.: Consistency of scoring rules: a reinvestigation of composition-consistency. Int. J. Game Theory **49**(3), 801–831 (2020). https://doi.org/10.1007/s00182-020-00711-7
17. Schneider, J., Schvartzman, A., Weinberg, S.M.: Condorcet-consistent and approximately strategyproof tournament rules. In: Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS) (2017)
18. Schvartzman, A., Weinberg, S.M., Zlatin, E., Zuo, A.: Approximately strategyproof tournament rules: on large manipulating sets and cover-consistence. In: 11th Innovations in Theoretical Computer Science Conference, ITCS 2020, 12–14 January 2020
19. Stanton, I., Williams, V.V.: Rigging tournament brackets for weaker players. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence, vol. 22, p. 357 (2011)