It's Not All Black and White: Degree of Truthfulness for Risk-Avoiding Agents

Eden Hartman, Erel Segal-Halevi, and Biaoshuai Tao

Abstract

The classic notion of *truthfulness* requires that no agent has a profitable manipulation — an untruthful report that, for *some* combination of reports of the other agents, increases her utility. This strong notion implicitly assumes that the manipulating agent either knows what all other agents are going to report, or is willing to take the risk and act as-if she knows their reports. Without knowledge of the others' reports, most manipulations are *risky* – they might decrease the manipulator's utility for some other combinations of reports by the other agents. Accordingly, a recent paper (Bu, Song and Tao, "On the existence of truthful fair cake cutting mechanisms", Artificial Intelligence 319 (2023), 103904) suggests a relaxed notion, which we refer to as *risk-avoiding truthfulness (RAT)*, which requires only that no agent can gain from a *safe* manipulation — one that is sometimes beneficial and never harmful.

Truthfulness and RAT are two extremes: the former considers manipulators with complete knowledge of others, whereas the latter considers manipulators with no knowledge at all. In reality, agents often know about some — but not all — of the other agents. This paper introduces the *RAT-degree* of a mechanism, defined as the smallest number of agents whose reports, if known, may allow another agent to safely manipulate, or n if there is no such number. This notion interpolates between classic truthfulness (degree n) and RAT (degree at least 1): a mechanism with a higher RAT-degree is harder to manipulate safely.

To illustrate the generality and applicability of this concept, we analyze the RAT-degree of prominent mechanisms across various social choice settings, including auctions, indivisible goods allocations, cake-cutting, voting, and two-sided matching.

Full Version: https://arxiv.org/abs/2502.18805

1 Introduction

The Holy Grail of mechanism design is the *truthful mechanism* — a mechanism in which the (weakly) dominant strategy of each agent is truthfully reporting her type. But in most settings, there is provably no truthful mechanism that satisfies other desirable properties such as budget-balance, efficiency or fairness. Practical mechanisms are thus *manipulable* in the sense that some agent a_i has a *profitable* manipulation – for *some* combination of reports by the other agents, agent a_i can induce the mechanism to yield an outcome (strictly) better for her by reporting non-truthfully.

This notion of manipulability implicitly assumes that the manipulating agent either knows the reports made by all other agents, or is willing to take the risk and act *as-if* she knows their reports. Without knowledge of the others' reports, most manipulations are *risky* – they might decrease the manipulator's utility for some other combinations of reports by the other agents. In practice, many agents are *risk-avoiding* and will not manipulate in such cases. This highlights a gap between the standard definition and the nature of such agents.

To illustrate, consider a simple example in the context of voting. Under the Plurality rule, agents vote for their favorite candidate, and the candidate with the most votes wins. If an agent knows that her preferred candidate has no chance of winning, she may find it beneficial to vote for her second-choice candidate to prevent an even less preferred candidate from winning. However, if the agent lacks precise knowledge of the other votes and decides to vote for her second choice, it may backfire – she might

inadvertently cause an outcome worse than if she had voted truthfully. For instance, if the other agents vote in a way that makes the agent the tie-breaker.

Indeed, various papers on cake-cutting (e.g. [13, 15]), voting (e.g. [46, 47, 31]) stable matching (e.g. Fernandez [26], Chen and Möller [18]) and coalition formation [54] studies truthfulness among such agents. In particular, Bu et al. [15] introduced a weaker notion of truthfulness, suitable for risk-avoiding agents, for cake-cutting. Their definition can be adapted to any problem as follows.

Let us first define a *safe manipulation* as a non-truthful report that may never harm the agent's utility. Based on that, a mechanism is *safely manipulable* if some agent a_i has a manipulation that is both profitable and safe; otherwise, the mechanism is Risk-Avoiding Truthful (RAT).

Standard truthfulness and RAT can be seen as two extremes with respect to safe-and-profitable manipulations: the former considers manipulators with complete knowledge of others, whereas the latter considers manipulators with no knowledge at all. In reality, agents often know about some — but not all — of the other agents.

This paper introduces the *RAT-Degree* — a new measurement that quantifies how robust a mechanism is to such safe-and-profitable manipulations. The RAT-Degree of a mechanism is an integer $d \in \{0, \ldots, n\}$, which represents — roughly — the smallest number of agents whose reports, if known, may allow another agent to safely manipulate; or n if there is no such number. (See Section 3 for formal definition).

This measure allows us to position mechanisms along a spectrum. A higher degree implies that an agent has to work harder in order to collect the information required for a successful manipulation; therefore it is less likely that mechanism will be manipulated. On one end of the spectrum are truthful mechanisms – where no agent can safely manipulate even with complete knowledge of all the other agents. The RAT-degree of such mechanisms is n. While on the other end are mechanisms that are safely manipulable – no knowledge about other agents is required to safely manipulate. The RAT-degree of such mechanisms is 0.

Importantly, the RAT-degree is determined by the worst-case scenario for the mechanism designer, which corresponds to the best-case scenario for the manipulating agents. The way we measure the amount of knowledge is based on a general objective applicable to all social choice settings.

Contributions. Our main contribution is the definition of the RAT-degree.

To illustrate the generality and usefulness of this concept, we selected several different social choice domains, and analyzed the RAT-degree of some prominent mechanisms in each domain. As our goal is mainly to illustrate the new concepts, we did not attempt to analyze all mechanisms and all special cases of each mechanism, but rather focused on some cases that allowed for a more simple analysis. To prove an upper bound on the RAT-degree, we need to show a profitable manipulation. However, in contrast to usual proofs of manipulability, we also have to analyze more carefully, how much knowledge on other agents is sufficient in order to guarantee the safety of the manipulation. This analysis gives more insight on the kind of manipulations possible in each mechanism, and on potential ways to avoid them. For clarity, we detail the results for each social choice setting in its corresponding section.

Organization. Section 2 introduces the model and required definitions. Section 3 presents the definition of RAT-degree. Section 4 explores auctions for a single good. Section 5 examines indivisible goods allocations. Section 5 focuses on cake cutting. Section 7 addresses single-winner ranked voting. Section 8 considers two-sided matching. Section 9 concludes with some future work directions.

Due to space constraints, this version provides only a summary of our results, together with intuition, main conclusions, and several open questions. Please refer to the extended version for more details.

1.1 Related Word

There is a vast body of work on truthfulness relaxations and alternative measurements of manipulability. Due to space constraints, we provide only a brief overview here; a more in-depth discussion of these works and their relation to RAT-degree can be found in Appendix B of the extended version.

Truthfulness Relaxations. Various truthfulness relaxations focus on a certain subset of all possible manipulations, which are considered more "likely". It requires that none of the manipulations from this subset is profitable. Different relaxations consider different subsets of "likely" manipulations. Brams et al. [13] propose *maximin strategy-proofness*, where an agent manipulates only if it is always beneficial. Waxman et al. [54] were the first (as far as we know) to use the term *safe manipulation*.[1] They examined the possible manipulations of agents with three different levels of knowledge on the social networks. Bu et al. [15] called a mechanism for cake-cutting, in which no agent has a safe-and-profitable manipulation, *risk-averse truthful (RAT)*; we prefer to call such a mechanism *risk-avoiding truthful*, as the definition assumes that agents completely avoid any risk. We extend their work by generalizing RAT to any social choice problem, and by suggesting a quantitative measure of the robustness of a mechanism to such manipulations. Troyan and Morrill [52] introduce *not-obvious manipulability (NOM)*, which assumes agents consider only extreme best or worst cases. RAT and NOM are independent notions. Fernandez [26] define *regret-free truth-telling (RFTT)*, where agents never regret truth-telling after observing the outcome. RAT and RFTT do not imply each other. Additionally, Slinko and White [46, 47], Hazon and Elkind [31] study "safe manipulations" in voting, but they consider coalition of voters and a different type of risk - that too many or too few participants will perform the exact safe manipulation. One can take a similar approach to measuring the degree with respect to most of these definitions. We believe this is an intriguing direction and leave it for future work.

Alternative Measurements. There are many approaches for quantifying manipulability from different perspectives. One approach considers the computational complexity of finding a profitable manipulation — e.g., [10, 11] (see [25, 53] for surveys). Another measurement is the number of bits an agent needs to know in order to have a safe manipulation, in the spirit of communication complexity, e.g., [40, 30, 14, 8] and compilation complexity - e.g., [21, 56, 33]. A third approach evaluates the probability that a profitable manipulation exists —e.g., [9, 35, 36]. The *incentive ratio*, which measures how much an agent can improve their utility by manipulating, is also widely studied—e.g., [16, 17, 38, 20, 19, 12, 51]. Other metrics include assessing the average and maximum gain per manipulation [2] and counting the number of agents who benefit from manipulating [4, 5]. Please refer to Appendix B of the extended version for a more thorough discussion of these alternative notions, and their relation to our notion.

2 Preliminaries

We consider a generic social choice setting, with a set of n *agents* $N = \{a_1, \ldots, a_n\}$, and a set of potential *outcomes* X. Each agent, $a_i \in N$, has preferences over the set of outcomes X, that can be described in one of two ways: (a) a linear ordering of the outcomes, or (b) a utility function from X to $\mathbb{R}$. The set of all possible preferences for agent a_i is denoted by $\mathcal{D}_i$, and is referred to as the agent's *domain*. We denote the agent's *true* preferences by $T_i \in \mathcal{D}_i$. Unless otherwise stated, when agent a_i weakly prefers the outcome x_1 over x_2, it is denoted by $x_1 \succeq_i x_2$; and when she strictly prefers x_1 over x_2, it is denoted by $x_1 \succ_i x_2$.

A *mechanism* or *rule* is a function $f : \mathcal{D}_1 \times \cdots \times \mathcal{D}_n \to X$, which takes as input a list of reported preferences $P_1, \ldots, P_n$ (which may differ from the true preferences), and returns the chosen outcome. In this paper, we focus on deterministic and single-valued mechanisms.

For any agent $a_i \in N$, we denote by $(P_i, \mathbf{P}_{-i})$ the preference profile in which agent a_i reports $P_i \in \mathcal{D}_i$,

[1] They use "safe manipulation" for a manipulation that is both safe and profitable.

and the other agents report $\mathbf{P}_{-i} \in \mathcal{D}_{-i}$ (where $\mathcal{D}_{-i} := \times_{j \in N \backslash \{i\}} \mathcal{D}_j$).

Truthfulness. A *manipulation* for a mechanism f and agent $a_i \in N$ is an untruthful report $P_i \in \mathcal{D}_i \backslash \{T_i\}$. A manipulation is *profitable* if there exists a set of preferences of the other agents for which it increases the manipulator's utility:

$$\exists \mathbf{P}_{-i} \in \mathcal{D}_{-i} : \ f(P_i, \mathbf{P}_{-i}) \succ_i f(T_i, \mathbf{P}_{-i}) \tag{1}$$

A mechanism f is called *manipulable* if some agent a_i has a profitable manipulation; otherwise f is called *truthful*.

RAT. A manipulation is *safe* if it never harms the manipulator's utility – it is weakly preferred over telling the truth for any possible preferences of the other agents:

$$\forall \mathbf{P}_{-i} \in \mathcal{D}_{-i} : \ f(P_i, \mathbf{P}_{-i}) \succeq_i f(T_i, \mathbf{P}_{-i}) \tag{2}$$

A mechanism f is called *safely-manipulable* if some agent a_i has a manipulations that is profitable and safe; otherwise f is called *risk-avoiding truthful (RAT)*.

3 The RAT-Degree

Let $a_i \in N$, $k \in \{0, \ldots, n-1\}$, $K \subseteq N \backslash \{a_i\}$ with $|K| = k$ and $\bar{K} := N \backslash (\{a_i\} \cup K)$. We denote by $(P_i, \mathbf{P}_K, \mathbf{P}_{\bar{K}})$ the preference profile in which the preferences of agent a_i are P_i, the preferences of the agents in K are $\mathbf{P}_K$, and the preferences of the agents in $\bar{K}$ are $\mathbf{P}_{\bar{K}}$.

Definition 1. Given an agent a_i, a subset $K \subseteq N \backslash \{a_i\}$ and preferences for them $\mathbf{P}_K \in \mathcal{D}_K$:

A manipulation P_i is called *profitable for a_i given K and $\mathbf{P}_K$* if

$$\exists \mathbf{P}_{\bar{K}} \in \mathcal{D}_{\bar{K}} : \ f(P_i, \mathbf{P}_K, \mathbf{P}_{\bar{K}}) \succ_i f(T_i, \mathbf{P}_K, \mathbf{P}_{\bar{K}}) \tag{3}$$

A manipulation P_i is called *safe for a_i given K and $\mathbf{P}_K$* if

$$\forall \mathbf{P}_{\bar{K}} \in \mathcal{D}_{\bar{K}} : \ f(P_i, \mathbf{P}_K, \mathbf{P}_{\bar{K}}) \succeq_i f(T_i, \mathbf{P}_K, \mathbf{P}_{\bar{K}}) \tag{4}$$

In words: The agents in K are those whose preferences are *Known* to a_i; the agents in $\bar{K}$ are those whose preferences are unknown to a_i. Given that the preferences of the known agents are $\mathbf{P}_K$, Equation (3) says that there exist a preference profile of the unknown agents that makes the manipulation profitable for agent a_i; while Equation (4) says that the manipulation is safe – it is weakly preferred over telling the truth for any preference profile of the unknown-agents.

The previous two definitions are special cases of Definition 1: Equation (1) - which defines a profitable manipulation - is equivalent to P_i being profitable given $\varnothing$; and Equation (2) - which defines a safe manipulation - is equivalent to P_i being safe given $\varnothing$.

Definition 2. A mechanism f is called *k-known-agents safely-manipulable* if for some agent a_i, some subset $K \subseteq N \backslash \{a_i\}$ with $|K| = k$ and some preferences for them $\mathbf{P}_K$, there exists a manipulation P_i that is both profitable and safe for a_i given K and $\mathbf{P}_K$.

Proposition 3.1. *Let $k \in \{0, \ldots, n-2\}$. If a mechanism is k-known-agents safely-manipulable, then it is also $(k+1)$-known-agents safely-manipulable.*

Proposition 3.1 justifies the following definition:

Definition 3. The *RAT-degree* of a mechanism f is the minimum k for which the mechanism is k-known-agent safely-manipulable, or n if there is no such k.

Intuitively, a mechanism with a higher RAT-degree is harder to manipulate, as a risk-avoiding agent would need to collect more information in order to find a safe manipulation.

Observation 3.2. *(a) A mechanism is truthful if-and-only-if its RAT-degree is* n.

(b) A mechanism is RAT if-and-only-if its RAT-degree is at least 1.

Figure 1 illustrates the relation between classes of different RAT-degree.

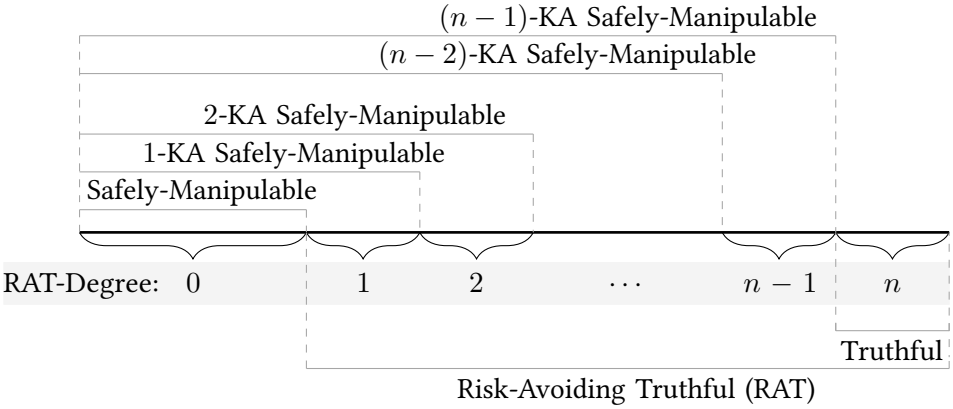

Figure 1: Hierarchy of the Manipulability and Truthfulness Classes with respect to the RAT-Degree. The horizontal axis represents the RAT-Degree, from 0 (safely-manipulable) to n (truthful). Labels above the axis correspond to Manipulability Classes, while labels below the axis correspond to Truthfulness Classes. KA stands for Known-Agents.

3.1 An Intuitive Point of View

Consider Table 1. We adopt the point of view of a particular agent a_i. The rows $T_i, P_i^1, \ldots, P_i^3$ correspond to the possible reports of agent a_i, where T_i is the truthful report and the rest are potential manipulations. The columns $\mathbf{P}_{-i}^1, \mathbf{P}_{-i}^2, \ldots$ represent possible strategy profiles of the remaining $n-1$ agents. The values $x_1, x_2 \ldots$ indicate the utility of agent a_i under each of these profiles when she reports truthfully.

When the risk-avoiding agent has no information (0-known-agents), a profitable-and-safe manipulation is a row in the table that represents an alternative report $P_i \neq T_i$, that (strictly) dominates T_i. That is, in each column, the outcome of the manipulation is at least as good as the truthful outcome, and in at least one column it is strictly better. In this example, P_i^1 satisfies this property.

When the risk-avoiding agent has more information (k-known-agents, when $k > 0$), it is equivalent to considering a strict subset of the columns. For instance, suppose the agent can infer—based on the information she has over some k other agents—that only the profiles $\mathbf{P}_{-i}^3, \mathbf{P}_{-i}^4, \mathbf{P}_{-i}^5$ are possible. Then P_i^2 is safe and profitable given this set. Notice that the utilities of agent a_i in profiles not among $\mathbf{P}_{-i}^3, \mathbf{P}_{-i}^4, \mathbf{P}_{-i}^5$ are irrelevant, since a_i considers them impossible.

Lastly, when the risk-avoiding agent has a full information ($(n-1)$-known-agents), she knows the exact strategy profile of the other agents, so it is equivalent to consider only one column in which the manipulation is profitable. P_i^3 in the table illustrates this type of manipulation.

4 Auction for a Single Good

We consider a seller owning a single good, and n potential buyers (the agents). The true preferences T_i of buyer a_i are given by real values $v_i \geqslant 0$, representing her happiness from receiving the good. The

	$\mathbf{P}^1_{-i}$	$\mathbf{P}^2_{-i}$	$\mathbf{P}^3_{-i}$	$\mathbf{P}^4_{-i}$	$\mathbf{P}^5_{-i}$	$\mathbf{P}^6_{-i}$	$\mathbf{P}^7_{-i}$	$\dots$
T_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	$\dots$
$P^1_i \neq T_i$	$\geqslant x_1$	$\geqslant x_2$	$\geqslant x_3$	$> x_4$	$\geqslant x_5$	$\geqslant x_6$	$\geqslant x_7$	$\dots$
$P^2_i \neq T_i$			$\geqslant x_3$	$> x_4$	$\geqslant x_5$			
$P^3_i \neq T_i$				$> x_4$				

Table 1: A Safe-And-Profitable Manipulation from an Agent Perspective. Dark-gray cells mean the value must be strictly higher, light-gray means it must be at least as high, and white means any value is allowed.

reported preferences P_i are the "bids" $b_i \geqslant 0$. A mechanism in this context has to determine the *winner* — the agent who will receive the good, and the *price* — how much the winner will pay. We assume that agents are quasi-linear – meaning their valuations can be interpreted in monetary units. Thus, the utility of the winning agent is the valuation minus the price; while the utility of the other agents is zero.

Results. Table 2 provides a summary of our results. The two most well-known mechanisms in this context are first-price auction and second-price auctions. First-price auction maximizes the seller's revenue when all buyers are truthful; but this assumption is, of course, unrealistic, as it is known to be manipulable. In fact, it is even safely-manipulable, so its RAT-degree is 0. We then prove that a first-price auction with a (fixed) positive discount has RAT-degree 1. Second-price auction is known to be truthful, so its RAT-degree is n. However, it has some important practical disadvantages [6]. In particular, when buyers are risk-averse, the expected revenue of a second-price auction is lower than that of a first-price auction [41]; even when the buyers are risk-neutral, a risk-averse *seller* would prefer the revenue distribution of a first-price auction [34].

This raises the question of whether it is possible to combine the advantages of both auction types. Indeed, we prove that any auction that applies a weighted average between the first-price and the second-price[2] achieves a RAT-degree of $n-1$, which is very close to being fully truthful (RAT-degree n). This implies that a manipulator agent would need to obtain information about all $n-1$ other agents to safely manipulate – which is a very challenging task. The seller's revenue from such an auction is higher compared to the second-price auction, giving this mechanism a significant advantage in this context. This result opens the door to exploring new mechanisms that are not truthful but come very close it. Such mechanisms may enable desirable properties that are unattainable with truthful mechanisms.

Mechanism	RAT-Degree
First-Price	0
First-Price With Positive Discount	1
Average Between First and Second Price	$n-1$
Second-Price	n

Table 2: Auctions for a Single Good: Summary of Results

5 Indivisible Goods Allocations

In this section, we consider several mechanisms to allocate m indivisible goods $G = \{g_1, \dots, g_m\}$ among the n agents. Here, the true preferences T_i of agent a_i are given by m real values: $v_{i,\ell} \geqslant 0$ for any $g_\ell \in G$, representing her happiness from receiving the good g_ℓ. The reported preferences P_i are real values $r_{i,\ell} \geqslant 0$. We assume the agents have *additive* valuations over the goods. Given a bundle

[2]The average price auction is mentioned as an exercise in [34] (Problem 2.4).

$S \subseteq G$, let $v_i(S) = \sum_{g_\ell \in S} v_{i,\ell}$ be agent a_i's utility upon receiving the bundle S. A mechanism in this context gets n (potentially untruthful) reports from all the agents and determines the *allocation* – a partition $(A_1, \ldots, A_n)$ of G, where A_i is the bundle received by agent a_i.

Results. Table 3 summarizes our results. We start by considering a simple mechanism, the *utilitarian* goods allocation – which assigns each good to the agent who reports the highest value for it. This mechanism always returns an allocation that maximizes the social welfare (i.e., sum of utilities) and is thus *Pareto Optimal* — there does not exist another allocation that all agents weakly prefer and at-least one strictly prefers. This mechanism is safely manipulable (RAT-degree = 0). We then show that the RAT-degree can be increased to 1 by requiring normalization — the values reported by each agent are scaled such that the set of all items has the same value for all agents.

We then consider the famous *round-robin* mechanism. It is easy to compute (specifically, polynomial) and guarantees *envy-freeness up to one good (EF1)* — each agent weakly prefers their own bundle over each other agent's bundle by the removal of at most one of the goods. Amanatidis et al. [3] prove that EF1 is incompatible with truthfulness for $m \geqslant 5$, which implies that the degree is at most $n - 1$. We prove that the situation is much worse, the RAT-degree of round-robin is at most 1. The proof relies on weak preferences (i.e., preferences that allow ties). In sharp contrast, we prove that when all agents have strict preferences, the RAT-degree is the best possible $n - 1$.

Lastly, we design a new mechanism that we call *volatile priority*. It satisfies EF1 and attains the best possible RAT-degree of $n - 1$ for all types of preferences; but does not run in polynomial time. The main observation behind our approach is that, in many cases, there are multiple solutions (here, allocations) that satisfy the required fairness properties (here, EF1). We use a priority order over the agents, determined by their reported valuations, to select among these solutions. Importantly, the priority ordering depends on *all* the agents' valuations to *all* the items. It is constructed such that, even a small change in one value, might induce a drastic change in the priority (see Section 5.4 in the extended version for details). As a result, a manipulating agent who lacks information about *all* other agents' valuations may unintentionally lower her priority by misreporting, and thus end up worse off. This makes manipulations risky. This raises the following open question regarding the computational complexity:

Open Question 5.1. *Is there a polynomial time EF1 goods allocation rule with RAT-degree $n - 1$?*

A natural candidate is The Envy-Cycle Elimination Mechanism by Lipton et al. [39].

Open Question 5.2. *What is the RAT-degree of The Envy-Cycle Elimination Mechanism?*

Mechanism		RAT-Degree	Properties
Utilitarian		0	Pareto Efficient
Normalized Utilitarian		1	Pareto Efficient (w.r.t. the normalized values)
Round-Robin (for $n > m$)	Any preferences	1	EF1, Polynomial-time
	Strict preferences	$n - 1$	
Volatile Priority*		$n - 1$	EF1

Table 3: Indivisible Goods Allocations: Summary of Results. (*) marks mechanisms introduced in this paper.

6 Cake Cutting

In this section, we study the *cake cutting* problem: the allocation of a divisible heterogeneous resource to n agents. The cake cutting problem was proposed by Steinhaus [48, 49], and it is a widely studied

subject in mathematics, computer science, economics, and political science.

In the cake cutting problem, the resource/cake is modeled as an interval $[0, 1]$, and it is to be allocated among a set of n agents $N = \{a_1, \ldots, a_n\}$. An allocation is denoted by $(A_1, \ldots, A_n)$ where $A_i \subseteq [0, 1]$ is the share allocated to agent a_i. We require that each A_i is a union of finitely many closed non-intersecting intervals, and, for each pair of $i, j \in [n]$, A_i and A_j can only intersect at interval endpoints, i.e., the measure of $A_i \cap A_j$ is 0.

The true preferences T_i of agent a_i are given by a *value density function* $v_i : [0, 1] \to \mathbb{R}_{\geqslant 0}$ that describes agent a_i's preference over the cake. To enable succinct encoding of the value density function, we adopt the widely considered assumption that each v_i is *piecewise constant*: there exist finitely many points $x_{i0}, x_{i1}, x_{i2}, \ldots, x_{ik_i}$ with $0 = x_{i0} < x_{i1} < x_{i2} < \cdots < x_{ik_i} = 1$ such that v_i is a constant on every interval $(x_{i\ell}, x_{i(\ell+1)})$, $\ell = 0, 1, \ldots, k_i - 1$. Given a subset $S \subseteq [0, 1]$ that is a union of finitely many closed non-intersecting intervals, agent a_i's value for receiving S is then given by $V_i(S) = \int_S v_i(x)dx$.

Lastly, we define an additional notion, *uniform segment*, which will be used throughout this section. Given n value density functions $v_1, \ldots, v_n$ (that are piecewise constant by our assumptions), we identify the set of points of discontinuity for each v_i and take the union of the n sets. Sorting these points by ascending order, we let $x_1, \ldots, x_{m-1}$ be all points of discontinuity for all the n value density functions. Let $x_0 = 0$ and $x_m = 1$. These points define k intervals, $(x_0, x_1), (x_1, x_2), \ldots, (x_{m-1}, x_m)$, such that each v_i is a constant on each of these intervals. We will call each of these intervals a *uniform segment*, and we will denote $X_t = (x_{t-1}, x_t)$ for each $t = 1, \ldots, m$.

Results. Our results are summarized in Table 4. We start by considering the *utilitarian* mechanism that outputs an allocation with the maximum social welfare. We show that the RAT-degree of this mechanism is 0. Similar as it is in the case of indivisible goods, we also consider the normalized variant of this mechanism, and we show that the RAT-degree is 1. Both mechanisms are *Pareto Efficient*, they always returns an allocation that is *Pareto Optimal* — there is no other allocation that all agents weakly prefer, and at least one agent strictly prefers.

We then consider several fair mechanisms that have been considered by Bu et al. [15]. Their paper studies whether or not these mechanisms are risk-avoiding truthful (in our language, whether the RAT-degree is positive). Here we provide a more fine-grained view.

Each mechanism is guaranteed to return an allocation satisfying at least one of the following properties: *Envy free (EF)* — each agent weakly prefers their own bundle to the bundle of any other agent, *Proportional* — each agent receives at least $1/n$ of the total value of the cake, according to their own valuation; and *Connected pieces* — each agent receives a single connected piece.

We start by considering *equal division mechanisms* — where we evenly allocate each uniform segment X_t to all agents. The output allocation is always envy-free and proportional. However, we show that the RAT-degree relies heavily on the order in which we allocate the pieces of each segment. Specifically, with fixed order (e.g., let agent a_1 get the left-most interval and agent a_n get the right-most interval), Bu et al. [15] already proved that it is not even RAT, which means that its RAT-degree is 0. To avoid this type of manipulation, a different tie-breaking rule was considered by Bu et al. [15] (See Mechanism 3 in their paper). We prove that its RAT-degree is $n - 1$. Tao [50] shows that *no* EF mechanism can also be truthful, so $n - 1$ is the best possible RAT-degree. However, this mechanism requires quite many cuts on the cake, and has a very poor performance in terms of pareto efficiency — it gives each agent a value of exactly $1/n$ and nothing more, which intuitively is the most inefficient allocation among the proportional ones.

We then focus on proportional mechanisms with connected pieces. Bu et al. [15] prove that the Dubins-Spanier's Moving-Knife [23] is not RAT, which in our terms means that its RAT-degree is 0. There was hope for those proven by Bu et al. [15] to have a positive RAT-degree: Bu, Song, and Tao's moving knife [15] (two variants - Mechanisms 4 and 5 in their paper); and Ortega and Segal-Halevi's moving knife

[42]. We show that they all have a very low RAT-degree of 1. This invoke the following question:

Open Question 6.1. *Is there a proportional connected cake-cutting rule with RAT-degree at least 2?*

A natural candidate to consider is the classic Even–Paz algorithm [24]; we conjecture that its RAT-degree is $\lceil n/2 \rceil$, but the question remains open.

Open Question 6.2. *What is the RAT-degree of the Even-Paz algorithm?*

Lastly, we adapt our *volatile priority* approach from the indivisible goods problem (Section 5), and propose a new mechanism that always returns a proportional and Pareto-efficient allocation, with the best possible RAT-degree of $n-1$. Unlike its counterpart, this mechanism runs in polynomial time. Given this result, it is natural to ask if the fairness guarantee can be strengthened to envy-freeness. A compelling candidate is the mechanism that always outputs allocations with maximum *Nash welfare* — the product of agents utilities. It is well-known that such an allocation is EF and Pareto-efficient. We conjecture the answer is $n-1$.

Open Question 6.3. *What is the RAT-degree of the maximum Nash welfare mechanism?*

Mechanism	RAT-Degree	Properties
Utilitarian	0	Pareto Efficient
Equal Division With Fixed Order	0	Proportional, EF
Dubins-Spanier's Moving-Knife [23]	0	Proportional, Connected
Normalized Utilitarian	1	Pareto Efficient (w.r.t. the normalized values)
Bu, Song, and Tao's Moving Knife [15] (Mechanism 4 and 5)	1	Proportional, Connected
Ortega and Segal-Halevi's Moving Knife [42]	1	Proportional, Connected
Bu, Song, and Tao's Equal Division [15] (Mechanism 3)	$n-1$	Proportional, EF
Volatile Priority*	$n-1$	Proportional, Pareto Efficient

Table 4: Cake Cutting: Summary of Results. (*) indicates that the mechanism was proposed in this paper. Gray cells indicate indicate results proven by Bu et al. [15].

7 Single-Winner Ranked Voting

We consider n voters (the agents) who need to elect one winner from a set C of m *candidates*. The agents' preferences are given by strict linear orderings $\succ_i$ over the candidates. When there are only two candidates, the majority rules and its variants (weighted majority rules) are truthful. With three or more candidates, the Gibbard–Satterthwaite Theorem [28, 44] implies that the only truthful rules are dictatorships. Our goal is to find non-dictatorial rules with a high RAT-degree.

We focus on *positional voting rules*, which parameterized by a vector of scores, $\mathbf{s} = (s_1, \ldots, s_m)$, where $s_1 \leqslant \cdots \leqslant s_m$ and $s_1 < s_m$. Each voter reports his entire ranking of the m candidates. Each such ranking is translated to an assignment of a score to each candidate: the lowest-ranked candidate is given a score of s_1, the second-lowest candidate is given s_2, etc., and the highest-ranked candidate is given a score of s_m. The total score of each candidate is the sum of scores he received from the rankings of all n voters. The winner is the candidate with the highest total score. If there are several

agents with the same maximum score, then the outcome is considered a tie. Common special cases of positional voting are *plurality voting*, in which $\mathbf{s} = (0, 0, 0, \ldots, 0, 1)$, and *anti-plurality voting*, in which $\mathbf{s} = (0, 1, 1, \ldots, 1, 1)$. By the Gibbard–Satterthwaite theorem, all positional voting rules are manipulable, so their RAT-degree is smaller than n. But, as we will show next, some positional rules have a higher RAT-degree than others.

Results. We show that all positional voting rules have an RAT-degree between $\approx n/m$ and $\approx n/2$. These bounds are almost tight: the upper bound is attained by plurality and the lower bound is attained by anti-plurality (up to small additive constants). These results raise the question of whether some other, non-positional voting rules have RAT-degrees substantially higher than $n/2$. Using our volatile priority approach (see Section 5), we could choose a "dictator", and take her first choice. This deterministic mechanism has RAT-degree $n-1$, any manipulation risks losing the chance to be the dictator. However, besides the fact that this is an unnatural mechanism, it suffers from other problems such as the *no-show paradox* (a participating voter might affect the selection rule in a way that will make another agent a dictator, which might be worse than not participating at all). Our main open problem is therefore to devise natural voting rules with a high RAT-degree.

Open Question 7.1. *Does there exist a non-dictatorial voting rule that satisfies the participation criterion (i.e. does not suffer from the no-show paradox), with RAT-degree larger than* $\lceil n/2 \rceil + 1$*?*

Mechanism		RAT-Degree	
		Lower Bound	Upper Bound
Positional Voting Rules	(assuming $n \geqslant 2m$)	$\lfloor (n+1)/m \rfloor - 1$	$\lceil n/2 \rceil + 1$
Plurality	(assuming $n \geqslant 5$)	$\lfloor n/2 \rfloor + 1$	$\lceil n/2 \rceil + 1$
Anti-Plurality	(assuming $n \geqslant m^2$)	$\lfloor (n+1)/m \rfloor - 1$	$\lfloor n/m \rfloor + 1$

Table 5: Single-Winner Ranked Voting with $m \geqslant 3$: Summary of Results.

8 Two-sided Matching

In this section, we consider mechanisms for two-sided matching. Here, the n agents are divided into two disjoint subsets, M and W, that need to be matched to each other. The most common examples are men and women or students and universities. Each agent has a strict preference order over the agents in the other set and being unmatched – for each $m \in M$, an order $\succ_m$ over $W \cup \{\phi\}$; and for each $w \in W$ an order, $\succ_w$, over $M \cup \{\phi\}$. A *matching* between M to W is a mapping μ from $M \cup W$ to $M \cup W \cup \{\phi\}$ such that (1) $\mu(m) \in W \cup \{\phi\}$ for each $m \in M$, (2) $\mu(w) \in M \cup \{\phi\}$ for each $w \in W$, and (3) $\mu(m) = w$ if and only if $\mu(w) = m$ for any $(m, w) \in M \times W$. When $\mu(a) = \phi$ it means that agent a is unmatched under μ. A mechanism in this context gets the preference orders of all agents and returns a matching. See Gonczarowski and Thomas [29] for a recent description of the structure of matching mechanisms.

Results. Our results for this problem are preliminary, so we provide only a brief overview here, with full descriptions and proofs in the appendix. We believe, however, that this is an important problem and that our new definition opens the door to many interesting questions.

We start with the well-known *deferred acceptance* mechanism of Gale and Shapley [27], which always returns a *stable* matching—no agent prefers being unmatched over their assigned match, and there is no pair of agents who would both prefer to be matched to each other over their current assignments. It is known that *no* stable matching mechanism is truthful for all agents [43]. Indeed, deferred acceptance is known to be truthful only for one side of the market. But what happens on the other side? Our

analysis reveals that the RAT-degree of deferred acceptance is very low: it is at least 1 and at most 3. The proof of the upper bound relies on *truncation*, where an agent falsely reports preferring to remain unmatched over certain options. We further show that even when agents are required to report complete preferences—thus ruling out this type of manipulation—the RAT-degree is at most 5. This raises the following important and interesting question:

Open Question 8.1. *Is there a stable matching mechanism with RAT-degree in* $\Omega(n)$*?*

We also examine the *Boston mechanism* [1], which is a widely used in practice for assigning students to schools. We establish an upper bound of 2 on its RAT-degree.

Mechanism	RAT-Degree		Properties
	Lower Bound	Upper Bound	
Deferred Acceptance (DA)	1	3	Stable, Truthful for M
DA under Complete Preferences	1	5	
Boston	0	2	

Table 6: Matchings: Summary of Results.

9 Discussion and Future Work

Our main goal in this paper is to encourage a more quantitative approach to truthfulness that can be applied to various problems. When truthfulness is incompatible with other desirable properties, we aim to find mechanisms that are "as hard to manipulate as possible", where hardness is measured by the amount of knowledge required for a safe manipulation.

Avoiding Risk. Our definition applies only to agents who entirely *avoid* even a small risk. This definition can be justified based on the well-known behavioral phenomenon called "zero-risk bias"[22, 55]. For example, experiments reported in the paper 'Prospect Theory: An Analysis of Decision under Risk', by Kahneman and Tversky [32], indicates that people underweight outcomes that are merely probable, in comparison with outcomes that are obtained with certainty. Another justification stems from our distinction between two types of risk: the risk due to the randomization of the mechanism, and the risk due to uncertainty about other agents' preferences. Our risk-avoidance assumption concerns only the second type of risk, for which there is usually no known probability distribution. Without such knowledge, agents are facing a non-quantifiable unknown risk (often referred to as *ambiguity*).[3] We believe that it is more reasonable to assume agents avoid this unknown risk altogether as they cannot even reliably distinguish between high and low risks. To compute expected utility with respect to this risk, one would need to adopt a Bayesian framework, which often requires strong and somewhat artificial assumptions about the strategies of other agents. We leave this direction for future work.

Randomized Mechanisms. This distinction between the two types of risk naturally leads to the topic of randomized mechanisms. In some settings, RAT-degree is more interesting for deterministic mechanisms, as many impossibility results apply only to deterministic mechanisms. However, in contexts such as voting, where there are impossibility results for randomized mechanisms too, the RAT-degree is applicable and potentially useful. In such settings, one can assume that agents avoid the non-quantifiable unknown risk that comes from the uncertainty about other agents' preferences, while treating the known risk induced by the mechanism construction using standard risk models—such as risk-neutral or risk-averse—aimed at maximizing expected utility.

Utility: Zero vs. Positive. Our model assumes that even a small positive gain can have a significant impact on behavior. For instance, the RAT-degree of the first-price auction is 0, but increases to 1

[3] This is related to another behavioral phenomenon known as "ambiguity aversion" or "uncertainty aversion".

when we introduce a fixed positive discount, even when the discount is arbitrarily small (see Section 4). This assumption is supported by findings in behavioral economics. For example, the paper 'Zero as a Special Price: The True Value of Free Products', by Shampanier et al. [45], shows that behavior changes significantly when something is free versus when it carries even a tiny cost. A related real-world example is the shift in behavior following the introduction of a small charge on plastic bags in many countries. When bags were free, most people used them without thinking; even a tiny positive cost was sufficient for 'nudging' people to bring reusable bags.

Tradeoffs. A particularly interesting future direction is to explore the tradeoffs between a high RAT-degree and other desirable properties, such as fairness. For example, in the case of two-sided matching (Section 8), we know that stability can be achieved with a low RAT-degree of at most 3, but is impossible to achieve with a RAT-degree of n. Where exactly does the boundary lie? Can we characterize the RAT-degree in relation to different fairness properties?

Beyond Worst-Case. We defined the RAT-degree as a "worst case" concept: to prove an upper bound, we find a single example of a safe manipulation. This is similar to the situation with the classic truthfulness notion, where to prove non-truthfulness, it is sufficient to find a single example of a manipulation. To go beyond the worst case, one could follow relaxations of truthfulness, such as truthful-in-expectation [37] or strategyproofness-in-the-large [7], and define similarly "RAT-degree in expectation" or "RAT-degree in the large".

Information About the Known-Agents. Our definition assumes that whenever an agent is known, we know their exact preference—that is, the precise preferences P_i they will report from their domain $\mathcal{D}_i$. In practice, however, we often have only partial information on the known-agents; for example, we might know that their preferences belong to a smaller subset of their domain. Consider Table 1 in Section 3.1. Such information still allows us to consider only a strict subset of the columns, leading to a weaker requirement than standard truthfulness. However, our results show that even under the assumption that we know the exact preferences, identifying the RAT-degree can already be highly nontrivial. For this reason, we only highlight this direction as a promising avenue for future research.

Changing Quantifiers. One could argue for a stronger definition requiring that a safe manipulation exists for every possible set of k known-agents, rather than for some set, or similarly for every possible preference profile for the known agents rather than just in some profile. However, we believe such definitions would be less informative, as in many cases, a manipulation that is possible for some set of k known-agents, is not possible for *any* such set. For example, in the first-price auction with discount (see Section 4), the RAT-degree is 1 under our definition. But if we required the the knowledge on any agent's bid would allow manipulation rather than just one, the degree would automatically jump to $n-1$, making the measure far less meaningful.

Combining the "known agents" concept with other notions. We believe that the "known agents" approach can be used to quantify the degree to which a mechanism is robust to other types of manipulations (besides safe manipulations), such as "always-profitable" manipulations or "obvious" manipulation. Accordingly, one can define the "max-min-strategyproofness degree" or the "NOM degree" (see Appendix B of the extended version).

Alternative Information Measurements. Another avenue for future work is to study other ways to quantify truthfulness. For example, instead of counting the number of known *agents*, one could count the number of *bits* that an agent should know about other agents' preferences in order to have a safe manipulation. The disadvantage of this approach is that different domains have different input formats, and therefore it would be hard to compare numbers of bits in different domains (see Appendix B of the extended version for more details).

Other applications. The RAT-degree can potentially be useful in any social-choice setting in which truthful mechanisms are not known, or lack other desirable properties. Examples include combinatorial auctions, multiwinner voting, budget aggregation and facility location.

10 Acknowledgement

This research is partly supported by the Israel Science Foundation grants 712/20, 3007/24, 2697/22 and 1092/24. The research of Biaoshuai Tao is supported by the National Natural Science Foundation of China (No. 62472271) and the Key Laboratory of Interdisciplinary Research of Computation and Economics (Shanghai University of Finance and Economics), Ministry of Education.

We sincerely appreciate Alexandros Psomas for his valuable contribution to this work. We are grateful to Sophie Bade, Avinatan Hassidim, Assaf Romm and Yonatan Aumann for their insights and helpful answers, to Paritosh Verma for his contribution; and to Michael Greinecker for his helpful responses.[4] Lastly, we would like to thank the reviewers in EC 2025 and COMSOC 2025 for their helpful comments.

References

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American economic review*, 93(3):729–747, 2003.

[2] Fuad Aleskerov and Eldeniz Kurbanov. Degree of manipulability of social choice procedures. *Current trends in Economics: theory and applications*, pages 13–27, 1999.

[3] Georgios Amanatidis, Georgios Birmpas, George Christodoulou, and Evangelos Markakis. Truthful allocation mechanisms without payments: Characterization and implications on fairness. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 545–562, 2017.

[4] Tommy Andersson, Lars Ehlers, and Lars-Gunnar Svensson. Budget balance, fairness, and minimal manipulability. *Theoretical Economics*, 9(3):753–777, 2014.

[5] Tommy Andersson, Lars Ehlers, and Lars-Gunnar Svensson. Least manipulable envy-free rules in economies with indivisibilities. *Mathematical Social Sciences*, 69:43–49, 2014.

[6] Lawrence M Ausubel and Paul Milgrom. The lovely but lonely vickrey auction. *Combinatorial auctions*, 17(3):22–26, 2006.

[7] Eduardo M Azevedo and Eric Budish. Strategy-proofness in the large. *The Review of Economic Studies*, 86(1):81–116, 2019.

[8] Yakov Babichenko, Shahar Dobzinski, and Noam Nisan. The communication complexity of local search. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 650–661, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367059. doi: 10.1145/3313276.3316354. URL https://doi.org/10.1145/3313276.3316354.

[9] Nathanaël Barrot, Jérôme Lang, and Makoto Yokoo. Manipulation of hamming-based approval voting for multiple referenda and committee elections. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS-2017)*, pages 597–605, 2017.

[10] John J Bartholdi, Craig A Tovey, and Michael A Trick. The computational difficulty of manipulating an election. *Social choice and welfare*, 6:227–241, 1989.

[11] John J Bartholdi III and James B Orlin. Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8(4):341–354, 1991.

[12] Xiaohui Bei, Biaoshuai Tao, Jiajun Wu, and Mingwei Yang. The incentive guarantees behind nash welfare in divisible resources allocation. *Artificial Intelligence*, page 104335, 2025.

[13] Steven J Brams, Michael A Jones, Christian Klamler, et al. Better ways to cut a cake. *Notices of the AMS*, 53(11):1314–1321, 2006.

[14] Simina Brânzei and Noam Nisan. Communication complexity of cake cutting. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, page 525, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367929. doi: 10.1145/3328526.3329644. URL https://doi.org/10.1145/3328526.3329644.

[15] Xiaolin Bu, Jiaxin Song, and Biaoshuai Tao. On existence of truthful fair cake cutting mech-

[4] https://economics.stackexchange.com/q/59899/385 and https://economics.stackexchange.com/a/24369/385

anisms. *Artificial Intelligence*, 319:103904, 2023. ISSN 0004-3702. doi: https://doi.org/10.1016/j.artint.2023.103904. URL https://www.sciencedirect.com/science/article/pii/S0004370223000504.

[16] Ning Chen, Xiaotie Deng, and Jie Zhang. How profitable are strategic behaviors in a market? In *Algorithms–ESA 2011: 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings 19*, pages 106–118. Springer, 2011.

[17] Ning Chen, Xiaotie Deng, Bo Tang, Hongyang Zhang, and Jie Zhang. Incentive ratio: A game theoretical analysis of market equilibria. *Information and Computation*, 285:104875, 2022.

[18] Yiqiu Chen and Markus Möller. Regret-free truth-telling in school choice with consent. *Theoretical Economics*, 19(2):635–666, 2024.

[19] Yu-Kun Cheng and Zi-Xin Zhou. An improved incentive ratio of the resource sharing on cycles. *Journal of the Operations Research Society of China*, 7:409–427, 2019.

[20] Yukun Cheng, Xiaotie Deng, Yuhao Li, and Xiang Yan. Tight incentive analysis on sybil attacks to market equilibrium of resource exchange over general networks. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 792–793, 2022.

[21] Yann Chevaleyre, Jérôme Lang, Nicolas Maudet, and Guillaume Ravilly-Abadie. Compiling the votes of a subelectorate. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-2009)*, pages 97–102, 2009.

[22] Daniel Crosby. *The Laws of Wealth*. Jaico Publishing House, 2021.

[23] Lester E Dubins and Edwin H Spanier. How to cut a cake fairly. *The American Mathematical Monthly*, 68(1P1):1–17, 1961.

[24] Shimon Even and Azaria Paz. A note on cake cutting. *Discrete Applied Mathematics*, 7(3):285–296, 1984.

[25] Piotr Faliszewski and Ariel D Procaccia. Ai's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64, 2010.

[26] Marcelo Ariel Fernandez. Deferred acceptance and regret-free truth-telling. Economics Working Paper Archive 65832, The Johns Hopkins University,Department of Economics, June 2018. URL https://ideas.repec.org/p/jhu/papers/65832.html.

[27] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[28] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601, 1973.

[29] Yannai A Gonczarowski and Clayton Thomas. Structural complexities of matching mechanisms. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 455–466, 2024.

[30] Elena Grigorieva, P Jean-Jacques Herings, Rudolf Müller, and Dries Vermeulen. The communication complexity of private value single-item auctions. *Operations Research Letters*, 34(5):491–498, 2006.

[31] Noam Hazon and Edith Elkind. Complexity of safe strategic voting. In *Algorithmic Game Theory: Third International Symposium, SAGT 2010, Athens, Greece, October 18-20, 2010. Proceedings 3*, pages 210–221. Springer, 2010.

[32] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.

[33] Neel Karia and Jérôme Lang. Compilation complexity of multi-winner voting rules (student abstract). In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, pages 15809–15810, 2021.

[34] Vijay Krishna. *Auction theory*. Academic press, 2009.

[35] Martin Lackner and Piotr Skowron. Approval-based multi-winner rules and strategic voting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-2018)*, pages 340–346, 2018.

[36] Martin Lackner, Jan Maly, and Oliviero Nardi. Free-riding in multi-issue decisions. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2023)*,

pages 2040–2048, 2023.
[37] Ron Lavi and Chaitanya Swamy. Truthful and near-optimal mechanism design via linear programming. *Journal of the ACM (JACM)*, 58(6):1–24, 2011.
[38] Bo Li, Ankang Sun, and Shiji Xing. Bounding the incentive ratio of the probabilistic serial rule. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, pages 1128–1136, 2024.
[39] Richard J Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. On approximately fair allocations of indivisible goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 125–131, 2004.
[40] Noam Nisan and Ilya Segal. The communication complexity of efficient allocation problems. *Draft. Second version March 5th*, pages 173–182, 2002.
[41] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.
[42] Josué Ortega and Erel Segal-Halevi. Obvious manipulations in cake-cutting. *Social Choice and Welfare*, 59(4):969–988, 2022.
[43] Alvin E Roth. The economics of matching: Stability and incentives. *Mathematics of operations research*, 7(4):617–628, 1982.
[44] 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.
[45] Kristina Shampanier, Nina Mazar, and Dan Ariely. Zero as a special price: The true value of free products. *Marketing science*, 26(6):742–757, 2007.
[46] Arkadii Slinko and Shaun White. Nondictatorial social choice rules are safely manipulable. In *Proceedings of the Second International Workshop on Computational Social Choice (COMSOC-2008)*, pages 403–414, 2008.
[47] Arkadii Slinko and Shaun White. Is it ever safe to vote strategically? *Social Choice and Welfare*, 43: 403–427, 2014.
[48] Hugo Steinhaus. The problem of fair division. *Econometrica*, 16(1):101–104, 1948.
[49] Hugo Steinhaus. Sur la division pragmatique. *Econometrica*, 17:315–319, 1949.
[50] Biaoshuai Tao. On existence of truthful fair cake cutting mechanisms. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 404–434, 2022.
[51] Biaoshuai Tao and Mingwei Yang. Fair and almost truthful mechanisms for additive valuations and beyond. In *International Conference on Web and Internet Economics*. Springer, 2024.
[52] Peter Troyan and Thayer Morrill. Obvious manipulations. *Journal of Economic Theory*, 185:104970, 2020.
[53] Yu A Veselova. Computational complexity of manipulation: a survey. *Automation and Remote Control*, 77:369–388, 2016.
[54] Naftali Waxman, Noam Hazon, and Sarit Kraus. Manipulation of k-coalitional games on social networks. *arXiv preprint arXiv:2105.09852*, 2021.
[55] Andrew Roman Wells and Kathy Williams Chiang. *Monetizing your data: A guide to turning data into profit-driving strategies and solutions*. John Wiley & Sons, 2017.
[56] Lirong Xia and Vincent Conitzer. Compilation complexity of common voting rules. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, pages 915–920, 2010.

Eden Hartman
Bar-Ilan University, Israel
Email: eden.r.hartman@gmail.com

Erel Segal-Halevi
Ariel University, Israel
Email: erelsgl@gmail.com

Biaoshuai Tao
Shanghai Jiao Tong University, China
Email: bstao@sjtu.edu.cn