Fair Allocations for Smoothed Utilities

YUSHI BAI, Tsinghua University
URIEL FEIGE, Weizmann Institute of Science
PAUL GÖLZ, Carnegie Mellon University
ARIEL D. PROCACCIA, Harvard University

When allocating indivisible items across agents, it is desirable for the allocation to be *envy-free*, which means that each agent prefers their own bundle over every other bundle. Even though envy-free allocations are not guaranteed to exist for worst-case utilities, they frequently exist in practice. To explain this phenomenon, prior work has shown that, if utilities are drawn from certain probability distributions, then envy-free allocations exist with high probability (as long as the number of items is sufficiently large relative to the number of agents).

In this paper, we study a more general setting, a *smoothed* model of utilities, in which utility profiles are mainly worst-case, but are slightly perturbed at random to avoid brittle counter-examples. Specifically, we start from a worst-case profile of utilities and, with some small probability, increase an agent's value for an item by adding a random amount, where the probability of perturbation and the distribution of perturbations are parameters of the model. If the probability of such perturbations is sufficiently large relative to the number of agents and items, we show that envy-free allocations exist with high probability and can be found efficiently. This analysis is tight up to constant factors. We also give an efficient algorithm for finding allocations that are simultaneously proportional and Pareto-optimal, which succeeds with high probability in the smoothed model.

CCS Concepts: • **Theory of computation → Algorithmic game theory and mechanism design**.

Additional Key Words and Phrases: fair division; envy freeness; smoothed analysis; semi-random models; beyond worst-case analysis

ACM Reference Format:
Yushi Bai, Uriel Feige, Paul Gölz, and Ariel D. Procaccia. 2022. Fair Allocations for Smoothed Utilities. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC '22), July 11–15, 2022, Boulder, CO, USA*. ACM, New York, NY, USA, 30 pages. https://doi.org/10.1145/3490486.3538285

EC '22, July 11–15, 2022, Boulder, CO, USA

ACM ISBN 978-1-4503-9150-4/22/07.
https://doi.org/10.1145/3490486.3538285

1 INTRODUCTION

Perhaps you are tasked with distributing a heap of candy between children, splitting an inheritance between siblings, or distributing the games included in a season ticket among your group of friends. Given that the recipients have different preferences over the items and that the items themselves cannot be individually split, can you distribute the items in a way that all recipients perceive as fair?

The most popular definition for whether an allocation of items is fair — some authors call it the "gold standard" [4, 22] — is a property known as *envy-freeness*: No agent should strictly prefer the bundle of another agent to their own. Multiple factors make envy-freeness an attractive notion of fairness. First, it avoids interpersonal comparisons of utility [2], which are contested from a philosophical angle [21]. Second, when utilities are additive — as we assume throughout this paper[1] — envy-freeness immediately implies other desirable notions of fairness such as proportionality and maximin share [12]. Finally, behavioral experiments have shown that avoiding envy is a factor in people's intuitive notion of fairness [25].

Given these advantages of envy-freeness, it is unfortunate that envy-free allocations do not always exist. An easy example is that of two agents and a single item that both agents value at a positive amount — no matter who receives the item, the agent who did not receive it will envy the agent who did. In light of such impossibilities, a large amount of recent work has focused on weaker variants of envy-freeness that *can* always be guaranteed such as *envy-freeness up to one good (EF1)* [6, 13, 14, 31].

Even though envy-free allocations do not exist in the worst case, they do exist in most real-world allocation problems. Indeed, among more than 4,300 tasks submitted through the "divide goods" application of the nonprofit website spliddit.org [23] as of July 2021, 71% have envy-free allocations. Moreover, the non-existence of envy-free allocations seems to primarily be a problem of instances with few items: Among the 401 instances with at least 3 times as many items as agents, 93% possess an envy-free allocation [36].

To explain why envy-free allocations frequently exist in practice, several papers [17, 33, 34] have studied conditions for the presence of envy-free allocations under the assumption that utilities are drawn from specific probability distributions. To illustrate this line of research, we focus on work by Dickerson, Goldman, Karp, Procaccia, and Sandholm [17]; we discuss related papers in Section 1.2. In the distributional model of Dickerson et al., envy-free allocations exist with high probability so long as the number of items exceeds the number of agents by a logarithmic factor. In fact, such envy-free allocation can be found by the simple *welfare-maximizing* algorithm, which allocates each item to the agent valuing it the most. It might be surprising that this algorithm — which, for general utilities, can be arbitrarily far from envy-free — would lead to envy-free allocations, but this follows from three assumptions underlying the distributional model: (1) for each item, each agent is equally likely to have the largest value, (2) when an agent has the largest value for a given item, this agent's conditional expected value for the item is larger by a constant gap than the expected value for the item conditioned on the agent's value not being the largest, and (3) agent–item values are independently drawn across items. Property (1) implies that each agent receives an equal number of items in expectation, property (2) implies that an agent's expected utility for an item in their own bundle is larger than their expected utility for an item in another bundle, and, by property (3), the agent's utilities for the different bundles are concentrated around their (well separated) means. Thus, with high probability, the welfare-maximizing allocation is envy-free.

[1] In practice, most fair division problems with indivisible goods assume additive utilities since additive utilities strike a good balance between expressivity and ease of elicitation. The website spliddit.org mentioned below and, to our knowledge, all papers studying fair allocations in distributional models assume additive utilities.

Even though papers such as the one by Dickerson et al. show that envy-free allocations are frequent when utilities are drawn from the distributions considered, the level of symmetry and independence in these distributions limits the applicability of these results to practice. For example, in the scenario of splitting season tickets among friends, the following two situations could naturally arise:

- Two friends largely agree on which games are the most exciting. However, one friend mainly wants to see the top matches whereas the other only slightly prefers these games over others.
- Some of the friends work at the same company. Since games on certain weekdays overlap with work commitments, these friends have zero utility for those games.

Whereas practical item allocation problems might exhibit many such patterns, these patterns are exceedingly unlikely in the distributional models. Since, furthermore, there is no reason to believe that algorithms like the welfare-maximizing algorithm would produce envy-free allocations on such instances, a stronger approach is needed to explain the widespread existence of envy-free allocations in practice.

There are two reasons to hope that weaker assumptions than those made in distributional models would suffice for envy-free allocations to exist: First, the argument by Dickerson et al. ensures not just that agents *barely* prefer their own bundle over another agent's, but that agents prefer their own bundle by a constant factor. Second, the research on EF1 shows that, even for worst-case utilities, it is always possible to allocate items such that no agent prefers another agent's bundle over their own by much, if no single item has exceptionally large value [31]. As a result, it stands to reason that utility profiles might not have to be entirely distributional to allow for envy-free allocations. Instead, utility profiles could be largely worst-case, with just a small part of the agents' utilities being distributional. As long as a subset of the items can be distributed in a suballocation in which agents strongly prefer their own bundle over others', and as long as the remaining items are distributed such that agents are only mildly envious of others' bundles in the second suballocation, the total allocation should be envy-free.

In this paper, we formalize these intuitions by proposing a *smoothed* model, in which utilities arise from a worst-case profile of utilities by randomly perturbing a fraction of the agent–item values. For any worst-case utility profile, envy-free allocations are likely to exist for the perturbed utility profile, if sufficiently many items are perturbed. Given that the perturbations only change a fraction of agent–item values and only change these values by a bounded amount, patterns in the worst-case utilities such as the ones described for the season-ticket scenario are largely preserved in the perturbed utilities and therefore such patterns do not rule out envy-free allocations. More fundamentally, our results show that instances without envy-free allocations are brittle with respect to small random variations in agent utilities (assuming a sufficient number of items), which provides a more robust explanation for the widespread existence of envy-free allocations than does previous work.

1.1 Our Techniques and Results

In Section 2, we introduce our smoothed model, which generates profiles of additive utilities for a given number n of agents and m indivisible items. The model starts from a profile of worst-case *base utilities*. With a fixed probability p, the model increases each agent–item base value by adding an independently drawn amount of utility, which we refer to as the agent's *boost* for that item. As we will discuss at the end of the paper, one can think of an item being boosted for an agent as the agent liking the item a bit more due to an *idiosyncratic* reason, i.e., a reason independent from general patterns of alignment or misalignment between the agents' base valuations. Both the boost probability and the distribution of boosts are parameters of our model, which allows our model to capture a wide range of distributions.

In Section 3, we demonstrate that the round robin algorithm, which has excellent envy-freeness properties in the distributional setting [34] and satisfies EF1 in the worst case, fails to guarantee envy-freeness in the smoothed model, even when the expected number $p\,m$ of boosts per agent is as large as $m^{0.33}$, which corresponds to a much larger p than will be required for our existence results.

In Section 4, we prove our main theorem:

Theorem 3 (informal statement)**.** *In the two following scenarios, envy-free allocations exist with probability converging to 1:*

- *if $n \to \infty$, $m \gg n \ln n$, and the expected number $p\,m$ of boosts per agent is sufficiently larger than $\ln n$, or*
- *if n is bounded and the expected number $p\,m$ of boosts per agent goes to infinity.*

Our proof applies an argument similar to the one by Dickerson et al. [17], but to only the boosts rather than the entire utilities: each item that has a boost for some agent is allocated to the agent with the largest boost to form a partial allocation $\mathcal{A}_2$; all items without any boosts are allocated using an algorithm satisfying EF1 in a suballocation $\mathcal{A}_1$.

Because the boosts satisfy assumptions (1) through (3) in the introduction, each agent derives substantially more utility from the boosts allocated to themselves in $\mathcal{A}_2$ than their utility for the boosts of items allocated to other agents in $\mathcal{A}_2$. While agents might have higher base utility for the items allocated to another agent than for those allocated to themselves, this source of envy is concentrated around zero (for $\mathcal{A}_2$) or bounded above by one (for $\mathcal{A}_1$). We show that the "negative envy" between boosts outweighs possible envy between base utilities, which makes the combination of $\mathcal{A}_1$ and $\mathcal{A}_2$ an envy-free allocation. In Section 5, we show that our analysis of the boost probability required for envy-free allocations to exist is asymptotically tight.

The argument for the existence of envy-free allocations sketched above does not provide a polynomial time allocation algorithm because the allocation assumed access to how agents' utilities decompose into base utilities and boosts, rather than to just the utilities. In Section 6, we develop a polynomial-time allocation algorithm for finding envy-free allocations with high probability, under the same conditions as those of Theorem 3. Our algorithm uses linear programming to identify a fractional suballocation that can play the role of $\mathcal{A}_2$, randomly rounds this allocation, and allocates the remaining items in an EF1 suballocation.

Finally, we show in Section 7 that our smoothed model makes another task easier that is computationally intractable in the worst case — finding allocations that simultaneously satisfy proportionality and Pareto-optimality. While the existence of such allocations is a corollary of Theorem 3, finding them in polynomial time requires additional techniques. Our algorithm is based on finding a fractional allocation using a different linear program, and rounding it using a different rounding procedure. The resulting allocation is always Pareto-optimal and satisfies proportionality with high probability over the random boosts.

1.2 Related Work

Fair allocations in distributional models. Our work is related to a sequence of papers that investigate the existence of fair allocations for indivisible items when utilities come from a distributional model [1, 5, 17, 18, 29, 32–34, 39, 42]. We already discussed in detail the paper by Dickerson et al. [17], and will discuss in the following two other papers whose results we will refer to.

These papers, both by Manurangsi and Suksompong [33, 34], refine Dickerson et al.'s analysis of parameter ranges for which envy-free allocations exist with high probability. Assuming that the number of agents n goes to infinity, Dickerson et al.'s results show that envy-free allocations exist if the number of items m scales in $\Omega(n \ln n)$, and that envy-free allocations do not necessarily exist if $m \in n + o(n)$. Manurangsi and Suksompong show that the existence for most intermediate m depends on how cleanly m is divisible by n, and that envy-free allocations cannot be guaranteed even if $m \in O(n \ln n / \ln \ln n)$ if m is not divisible by n [33]. In a follow-up paper [34], Manurangsi and Suksompong show that another algorithm, round robin, leads to envy-free allocations for $m \in \Omega(n \ln n / \ln \ln n)$, rendering the previously mentioned bound tight up to constant factors. In our smoothed model, by contrast, the round robin algorithm will turn out to be far from optimal for guaranteeing envy-freeness.

Semi-random models. Smoothed analysis is one modeling approach within the wider field of *semi-random models*. Semi-random models generate problem instances in a partly adversarial, partly random way. These semi-random models allow algorithm designers to study algorithmic problems despite impossibility results holding for isolated worst-case instances, without forcing the algorithm designer to commit to any specific distribution as being a good proxy for reality.

While semi-random models have been widely adopted as a way to circumvent computational hardness [10, 19, 20, 28, 37], they have only rarely been used to circumvent other kinds of impossibilities as we do in our paper. One such example is in online matching, where worst-case arrivals with random ordering allow for better approximation ratios than worst-case arrival orderings [24, 26, 27]. Recently, semi-random models have also been used in social choice [40, 41], where the randomness in the instance circumvents classic voting paradoxes, and in mechanism design [8, 9], where the randomness allows to align incentives without the large cost to welfare implied by worst-case examples. The only other application of semi-random models to fair division that we know of is by Chèze [15]. He applies smoothed analysis to the cake-cutting problem, in which envy-free allocations always exist but are hard to compute. For smoothed utilities, Chèze shows that the complexity of cake cutting is polynomially bounded with high probability.

2 PRELIMINARIES

Basic definitions. Let there be a population $N = \{1, \dots, n\}$ of $n \geq 2$ agents and a collection M of $m \geq 1$ indivisible items. Each agent i has a *value* $u_i(\alpha) \geq 0$ for each item α. Throughout this paper, we assume that the agents' utilities are additive, i.e., that agent i's utility for a set of items $A \subseteq M$ is $u_i(A) := \sum_{\alpha \in A} u_i(\alpha)$. The collection of all agent–item values is called a *utility profile*. Given a utility profile, an allocation algorithm must return an *allocation*, which is a partition of the items into *bundles* $A_1, \dots, A_n$ where bundle A_i is allocated to agent i.

Given an allocation $\{A_i\}_{i \in N}$ for a specific utility profile, we say that an agent i *envies* another agent j by the amount $u_i(A_j) - u_i(A_i)$. Note that we allow envy to be negative, whereas some authors define envy as $\max\big(0, u_i(A_j) - u_i(A_i)\big)$. We say that an allocation has t-bounded envy for some $t \in \mathbb{R}$ if each agent i envies j by at most t; if an allocation has t-bounded envy for $t < 0$, this means that each agent strictly prefers their own bundle over every other bundle by a gap of size $-t$. An allocation $\{A_i\}_{i \in N}$ is *envy-free* if it has 0-bounded envy, i.e., if, for all pairs of agents $i \neq j$, $u_i(A_i) \geq u_i(A_j)$. Rather than defining *envy-freeness up to one good* (EF1), it will suffice to note that, if an allocation satisfies EF1, the allocation has $\max_{i \in N, \alpha \in M} u_i(\alpha)$-bounded envy. An allocation $\{A_i\}_{i \in N}$ satisfies *proportionality* if, for every agent i, $u_i(A_i) \geq u_i(M)/n$. Note that, for additive valuations, any envy-free allocation is also proportional. Finally, an allocation $\{A_i\}_{i \in N}$ is *Pareto-optimal* if it is not *Pareto dominated* by any other allocation $\{A'_i\}_{i \in N}$, which means that there can be no $\{A'_i\}_{i \in N}$ such that, for each agent i, $u_i(A'_i) \geq u_i(A_i)$ and at least one of these inequalities is strict.

Smoothed model. Each *instance* of our smoothed model describes a family of probability distributions that generate utility profiles. These instances are parameterized by three elements: a utility profile of *base utilities*, a *boost probability*, and a *boost distribution*: Each agent i has a *base value* $\underline{u}_i(\alpha) \in [0, 1]$ for each item α, and the *base utilities* extend these values to sets of items by setting $\underline{u}_i(A) \coloneqq \sum_{\alpha \in A} \underline{u}_i(\alpha)$. Independently across agents and items, an agent i's *boost* for item α, denoted $Boost_{i,\alpha}$, is drawn from a *boost distribution* $\mathcal{D}$ with a fixed *boost probability* $p \in (0, 1]$; with probability $1 - p$, the boost is zero.[2] We assume that all values drawn from $\mathcal{D}$ lie in the interval $(0, c]$ for some $c > 0$ and that $\mathcal{D}$ is non-atomic.[3] We call $\mathcal{D}$'s variance σ^2, which is positive by non-atomicity. An agent's value for an item α in the smoothed model is defined as $u_i(\alpha) \coloneqq \underline{u}_i(\alpha) + Boost_{i,\alpha}$.

In this paper, we aim to characterize how many random boosts are needed in order to have envy-free allocations, and we focus on the case where there are many items. We phrase these results in terms of asymptotic bounds on the boost probability p, along an infinite sequence of instances in which the number of items goes to infinity. For an event relating to a utility profile (such as: "there exists an envy-free allocation"), we say that this event occurs *with high probability* if, along the infinite sequence of instances, the probability of this event converges to 1. For simplicity, we will assume that the boost distribution $\mathcal{D}$ stays the same across the sequence of instances instances (alternatively, that $\sigma^2 > 0$ refers to a lower bound on all variances and c to an upper bound on the support of all boost distributions). Note that we do not necessarily assume that the number of agents n goes to infinity; we will often give separate results depending on whether $n \to \infty$ or $n \in \mathcal{O}(1)$. For clarity, we use the notation $const(\sigma^2, c)$ to hide positive, absolute constants that only depend on the variance and upper range of $\mathcal{D}$.

In Appendix A, we give the rationale for three central modeling choices underlying our smoothed model. Specifically, we explain why our model assumes boosts to be positive, why it includes the parameter p, and why we focus on the case where m goes to infinity. For each of these decisions, we discuss the implications of alternative choices. In addition, we describe in Appendix G how our results extend to a model of allocating not just goods but mixed manna (i.e., items whose values may be negative), a setting in which we make the first and second modeling choice differently.

3 ROUND ROBIN ALLOCATIONS ARE NOT ENVY-FREE

Given that our smoothed model generalizes the distributional model mentioned in the introduction,[4] we will demonstrate that the allocation algorithms shown to give envy-freeness in this distributional model do not guarantee envy-freeness in the smoothed model even when the number of boosts is high. This finding will show that the smoothed setting requires new algorithms such as those we develop in the following sections. The welfare-maximizing algorithm is clearly not useful for finding envy-free allocations in this model since, if one agent's base utilities exceed another agent's base utilities by more than c for all items, the welfare-maximizing algorithm will not allocate any item to the latter agent, who will then be envious. Thus, we will focus our discussion on another candidate algorithm.

A priori, the most promising contender for an algorithm that would produce envy-free allocations in the smoothed model is the *round robin algorithm*. This common-sense algorithm iterates over

[2] Our results are robust to some changes in this setup. For example, boost distributions could differ by agent, in which case our analysis would extend using the tools of Kurokawa et al. [29, Lemma 3.2] or Bai and Gölz [5]. Fundamentally, all we need is for the boost distribution to satisfy properties (1), (2), and (3) from the introduction.

[3] The assumption of non-atomicity is only for ease of exposition; our results readily generalize to atomic distributions with positive variance when ties between agents' boosts are broken at random in our arguments.

[4] That is, if $p = 1$ and all base utilities are zero, the smoothed model reduces to the distributional model in which all agent–item utilities are drawn i.i.d. from a distribution $\mathcal{D}$.

the agents i in a cyclical manner, and allocates one item to i in each step, namely, i's favorite item among those that have not yet been allocated. Round robin excels by virtue of its envy properties both in the distributional model and in the worst case: In the distributional model, round robin allocations are envy-free with high probability if $m \in \Omega(n \ln n / \ln \ln n)$ [34], which is the best possible such guarantee since envy-free allocations need not exist for fewer items [33]. Even on worst-case valuations, round robin satisfies EF1 [13].

On the face of it, the argument showing EF1 in the worst case seems like it should also provide envy-freeness for smoothed utilities. To describe this argument for EF1, consider two agents 1 and 2 and suppose that they draw items as follows, possibly interspersed by the picks of other agents: agent 1 picks an item α_1, agent 2 picks β_1, agent 1 picks α_2, agent 2 picks β_2, and so forth. By construction, agent 1 prefers each item α_t to agent 2's subsequent pick β_t, from which it follows that $u_1(\{\alpha_1, \alpha_2, \dots\}) \geq u_1(\{\beta_1, \beta_2, \dots\})$. Agent 1 can only envy agent 2 if agent 2 picked one additional item β_0 before agent 1 chose an item for the first time, but even then the envy is bounded by the value of β_0, which shows EF1.

It would be natural to assume that few random utility perturbations suffice in order to eliminate the envy in round robin allocations, since envy created by β_0 persists only if the $u_1(\alpha_t)$ are not much larger than the corresponding $u_1(\beta_t)$, which will require the two agents' preferences to closely align. Random perturbations of the utilities seem promising because they disturb the alignment of preferences even if the base utilities are similar; agent 1's boosts will make some items particularly valuable to agent 1 and some boosts will make agent 2 pick items earlier that have low value for agent 1. If these effects make $u_1(\alpha_t) - u_1(\beta_t)$ large enough for enough t, then agent 1 prefers the items $\alpha_1, \alpha_2, \dots$ by a sufficient gap over $\beta_1, \beta_2, \dots$ that adding β_0 to the latter bundle does not create envy. Surprisingly to us, the following proposition shows that even a substantial amount of random boosts is not enough to eliminate envy for certain base utilities:

Proposition 1. *For any $\delta > 0$ and $\epsilon \in (0, 1)$, there is a family of instances with two agents, 1 and 2, and $m \to \infty$ items, with a boost probability of $p := m^{-2/3-\delta}$ and the boost distribution Uniform$\big((0, \epsilon)\big)$ such that, with high probability, agent 1 will envy agent 2 in the round robin allocation starting from agent 2.*

Proof sketch. We describe the family of instances and their non-trivial analysis in Appendix B. The idea behind these instances is to set the base utilities such that the boosts do not have the envy-reducing effects for agent 1 described above: Whenever agent 1 picks an item that is boosted for them, this gives agent 2 an opportunity to pick more items that both agents have high value for, essentially neutralizing any envy-reducing effects of agent 1's boosts. Furthermore, the instance is constructed such that agent 2's boosts only change their sequence of picking between items that agent 1 is indifferent between, and thus agent 2's boosts do not reduce agent 1's envy. As a result, agent 1 envies agent 2, even though the expected number of boosted items $p\,m$ grows as a power function of m. □

In contrast to this negative result for round robin, we will show in the next section that envy-free allocations exist for two agents whenever $p\,m \to \infty$ and thus already for boost probabilities much lower than those in this proof. Note that the proof above also rules out proportional allocations since proportionality and envy-freeness are equivalent when $n = 2$. One might hope that the counter-example above is a fluke of having only two agents, but we show in Appendix B that the argument can easily accommodate additional agents (who favor the items that agent 1 and 2 value the least):

Corollary 2. *Fix any $\delta > 0$, $\epsilon \in (0, 1)$ and any function $n : \mathbb{N}_{\geq 1} \to \mathbb{N}_{\geq 2}$ such that $n(m) < m^{1/3}$ for sufficiently large m. Then, there is a family of instances with $m \to \infty$ items and $n = n(m)$ agents, with* *a boost probability of $p := m^{-2/3-\delta}$ and the boost distribution Uniform$\big((0,\epsilon)\big)$ such that, with high probability, round robin (starting with agent 2) produces an allocation in which agent 1 envies agent 2.*

4 EXISTENCE OF ENVY-FREE ALLOCATIONS

We now prove that envy-free allocations are likely to exist, implementing the argument sketched in Section 1.1. Formally, consider a procedure[5] that allocates the items in two phases: In the first phase, it allocates all items that are not boosted for any agent via an EF1 procedure such as round robin. Let A_i^1 denote the set of items received by agent i in phase 1. In the second phase, the procedure gives each remaining item to the agent with the highest boost.[6] Let A_i^2 denote the set of items received by agent i in phase 2, and let $A_i := A_i^1 \dot\cup A_i^2$ denote agent i's total allocation. Our main theorem states that this allocation is likely to be envy-free:

Theorem 3. *There exists a constant $K_{\sigma^2,c} > 0$ (that only depends on σ^2 and c), such that,*

- *if $n \to \infty$ and $m \to \infty$, $m \geq const(\sigma^2, c)\, n \ln n$, and $p \geq \frac{const(\sigma^2,c)\,\ln n}{m}$, or*
- *if $m \to \infty$, n is bounded, and $p \in \omega(1/m)$,*

then, with high probability, there is an allocation with $\big(1 - K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m\big)$-bounded envy. In particular, under the conditions of the theorem, there exists an envy-free allocation with high probability.

In the remainder of this section, we prove this theorem: In Section 4.1, we prove that each agent i has higher utility for the boosts in A_i^2 than for the boosts in any other A_j^2. Next, in Section 4.2, we bound the amount by which an agent i's base utility for a bundle A_j^2 can exceed i's base utility for A_i^2. In Section 4.3, we combine the previous two lemmas with the EF1 property of $\{A_i^1\}_{i \in N}$ to prove the main theorem.

4.1 Negative Envy of Boosts

We begin by analyzing the allocation of the boosts. Specifically, we will show that, with high probability, an agent i's bundle contains a larger amount of boosts for agent i than the bundles given to other agents. While the structure of our proof resembles the argument by Dickerson et al. [17], proving the linear gap between the expected utility of an agent's own bundle and the expected utility of another agent's bundle is a bit more difficult. The main difference stems from the fact that, in the model of Dickerson et al., each item's value for the item's recipient is the largest out of n random utilities. In our setting, by contrast, only a random subset of agents have a boost drawn from $\mathcal{D}$. Thus, a careful bound is required to show that (for agents $i \neq j$ and an item α) $\mathbb{E}[u_i(\alpha) \mid \alpha \in A_i^2] - \mathbb{E}[u_i(\alpha) \mid \alpha \in A_j^2]$ is lower-bounded by a positive constant $C_{\sigma^2,c}$ that is independent of p and of $\mathcal{D}$'s properties other than σ^2 and c.

In our analysis, a very useful parameter will be the probability r that a given item α will be part of a given agent i's phase-2 allocation. By symmetry, this probability is

$$r := \mathbb{P}[\alpha \in A_i^2] = \frac{\mathbb{P}[\alpha \text{ is boosted for some agent}]}{n} = \frac{1-(1-p)^n}{n}.$$

When p is small, r is approximately equal to p, but once p grows larger than $1/n$, r grows slower and slower in p and instead converges towards $1/n$. In fact, we have that $(1 - 1/e)\min(p, 1/n) \leq r \leq \min(p, 1/n)$, as we prove in Lemma 11 in Appendix C.

[5] We avoid talking about an algorithm here, for reasons laid out in Section 6.

[6] We will ignore the possibility of ties between positive boosts, which, since $\mathcal{D}$ is non-atomic, occur with zero probability.

Lemma 4. *There exists a constant $K_{\sigma^2,c} > 0$ that depends only on σ^2 and c such that, with probability at least $1 - n^2\, \exp\left(-const\,(\sigma^2, c)\, r\, m\right)$, for all agents $i \neq j$, it holds that*

$$\sum_{\alpha \in A_i^2} Boost_{i,\alpha} - \sum_{\alpha \in A_j^2} Boost_{i,\alpha} \geq 6\, K_{\sigma^2,c}\, r\, m. \tag{1}$$

Proof. Fix an item α and an agent i. To analyze the item's boosts, we use the principle of deferred decision by generating the boosts of the item α in two stages, as follows:

In the first stage, we only determine which of the following (exhaustive and mutually exclusive) cases applies for the boosts of item α:

Case	Case description	Who receives α?	Probability
I	$Boost_{i,\alpha} = 0$	depends	$1 - p$
II	$Boost_{i,\alpha} > 0$ and $\forall j \neq i.\ Boost_{j,\alpha} = 0$	i	$p\,(1-p)^{n-1}$
III	$Boost_{i,\alpha} = \max_{j \in N} Boost_{j,\alpha} > 0$ and $\exists j \neq i.\ Boost_{j,\alpha} > 0$	i	$r - p\,(1-p)^{n-1}$
IV_j for $j \neq i$	$Boost_{i,\alpha} > 0$ and $Boost_{j,\alpha} = \max_{j' \in N} Boost_{j',\alpha} > 0$	j	$(p-r)/(n-1)$

In the second stage, we sample i's boost for α conditioned on the case determined in the first stage. We will use the following observations about these conditional distributions:

Case I: By definition, $Boost_{i,\alpha} = 0$.

Case II: Since the boosts of other agents are independent, we have learned that $Boost_{i,\alpha}$ is positive, but nothing else about it. Thus, the conditional distribution is $\mathcal{D}$, with expected value $\mathbb{E}[\mathcal{D}]$.

Case III: The fact that $Boost_{i,\alpha}$ is the greatest out of $k \geq 2$ independent draws means that the conditional distribution is the kth order statistic of k independent draws from $\mathcal{D}$. For any k, this distribution stochastically dominates the second order statistic of two independent draws, which we will denote by $\mathcal{D}_{2:2}$. The expectation is at least $\mathbb{E}[\mathcal{D}_{2:2}]$.

Cases IV_j. The information that, out of $k \geq 2$ independent samples from $\mathcal{D}$, another draw is larger than $Boost_{i,\alpha}$ implies that its conditional distribution is stochastically dominated by $\mathcal{D}$. Thus, the conditional expectation is at most $\mathbb{E}[\mathcal{D}]$.

By the above, for each item α, we can bound the expected contribution of its boosts to the sums in Eq. (1). We denote α's expected contribution to $\sum_{\alpha \in A_i^2} Boost_{i,\alpha}$ by μ_{own} and denote α's expected contribution to $\sum_{\alpha \in A_j^2} Boost_{i,\alpha}$ for an arbitrary agent $j \neq i$ by μ_{other}:

$$\begin{aligned}
\mu_{\text{own}} &:= \mathbb{E}\left[\mathbb{1}_{\alpha \in A_i^2} \cdot Boost_{i,\alpha}\right] \\
&= \mathbb{P}[\alpha \text{ in case II}] \cdot \mathbb{E}[Boost_{i,\alpha} \mid \alpha \text{ in case II}] + \mathbb{P}[\alpha \text{ in case III}] \cdot \mathbb{E}[Boost_{i,\alpha} \mid \alpha \text{ in case III}] \\
&\geq p\,(1-p)^{n-1}\, \mathbb{E}[\mathcal{D}] + (r - p\,(1-p)^{n-1})\, \mathbb{E}[\mathcal{D}_{2:2}], \text{ and} \\
\mu_{\text{other}} &:= \mathbb{E}\left[\mathbb{1}_{\alpha \in A_j^2} \cdot Boost_{i,\alpha}\right] \\
&= \mathbb{P}[\alpha \text{ in case I and } \alpha \in A_j^2] \cdot 0 + \mathbb{P}[\alpha \text{ in case IV}_j] \cdot \mathbb{E}[Boost_{i,\alpha} \mid \alpha \text{ in case IV}_j] \\
&\leq \frac{p-r}{n-1}\, \mathbb{E}[\mathcal{D}].
\end{aligned}$$

Next, we observe that $r - p\,(1-p)^{n-1} \geq (p-r)/(n-1)$. The left-hand side describes the probability (r) that agent i has the largest positive boost, but i is not the only agent to have a positive boost (we subtract the probability $p\,(1-p)^{n-1}$ with which only i has positive boost). Hence the left hand side gives the probability that i has largest positive boost, and in addition to i there is at least one agent $j \neq i$ with positive boost. The numerator of right-hand side describes the probability that a fixed agent $j \neq i$ satisfies two conditions: it has a positive boost (with probability p) but not the largest boost (hence we subtract r). As each agent other than j is equally likely to have the largest boost, dividing by $n-1$ gives the probability that i has the largest boost. Hence the right hand side gives the probability that i has largest positive boost, and in addition to i some specific agent $j \neq i$ has positive boost. Because the right hand side event implies the left hand side event, the inequality holds. This observation allows us to lower-bound the gap Δ_μ as follows:

$$\begin{aligned}
\Delta_\mu &\coloneqq \mu_{\text{own}} - \mu_{\text{other}} \\
&\geq p\,(1-p)^{n-1}\,\mathbb{E}[\mathcal{D}] + (r - p\,(1-p)^{n-1})\,\mathbb{E}[\mathcal{D}_{2:2}] - \frac{p-r}{n-1}\,\mathbb{E}[\mathcal{D}] \\
&\geq p\,(1-p)^{n-1}\,\mathbb{E}[\mathcal{D}] + (r - p\,(1-p)^{n-1})\,\mathbb{E}[\mathcal{D}_{2:2}] - (r - p\,(1-p)^{n-1})\,\mathbb{E}[\mathcal{D}] \\
&= p\,(1-p)^{n-1}\,\mathbb{E}[\mathcal{D}] + (r - p\,(1-p)^{n-1})\,(\mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}]) \\
&\geq r \cdot \min(\mathbb{E}[\mathcal{D}], \mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}]) \geq C_{\sigma^2,c}\, r,
\end{aligned}$$

where the last inequality is from the fact that both $\mathbb{E}[\mathcal{D}]$ and $\mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}]$ are lower-bounded by a constant $C_{\sigma^2,c} > 0$ that only depends on the variance of $\mathcal{D}$ and on the upper end of its range c. This is implied by Lemma 13 in Appendix C.

Finally, note that we can bound $\mu_{\text{own}} = \mathbb{E}\left[\mathbb{1}_{\alpha \in A_i^2} \cdot \mathit{Boost}_{i,\alpha}\right] \leq \mathbb{E}\left[\mathbb{1}_{\alpha \in A_i^2} \cdot c\right] \leq r\,c$ and, analogously, $\mu_{\text{other}} \leq r\,c$.

We will now lower-bound agent i's sum of boost for the items that were allocated to them, i.e., the sum $\sum_{\alpha \in A_i^2} \mathit{Boost}_{i,\alpha} = \sum_{\alpha \in M} \mathbb{1}_{\alpha \in A_i^2} \cdot \mathit{Boost}_{i,\alpha}$. Recall that all boosts are in $(0, c]$ and are independent across items α, which allows us to apply a Chernoff bound:

$$\begin{aligned}
&\mathbb{P}\left[\sum_{\alpha \in A_i^2} \mathit{Boost}_{i,\alpha} \leq (\mu_{\text{own}} - \Delta_\mu/3)\, m\right] = \mathbb{P}\left[\sum_{\alpha \in A_i^2} \mathit{Boost}_{i,\alpha} \leq \left(1 - \frac{\Delta_\mu}{3\,\mu_{\text{own}}}\right) \mu_{\text{own}}\, m\right] \\
&\leq \exp\left(-\frac{\left(\frac{\Delta_\mu}{3\,\mu_{\text{own}}}\right)^2 \mu_{\text{own}}\, m}{2\,c}\right) = \exp\left(-\frac{\Delta_\mu^2\, m}{18\,c\,\mu_{\text{own}}}\right) \leq \exp\left(-\frac{C_{\sigma^2,c}^2\, r^2\, m}{18\,c^2\, r}\right) = \exp\left(-\frac{C_{\sigma^2,c}^2}{18\,c^2}\, r\, m\right).
\end{aligned}$$

Next, we will upper-bound the sum of i's boosts for all items allocated to a fixed agent $j \neq i$, i.e., the sum $\sum_{\alpha \in A_j^2} \mathit{Boost}_{i,\alpha}$. For this, it suffices to consider items α of case IV$_j$; items α of case I might be in A_j^2, but contribute zero to the sum. We bound

$$\begin{aligned}
&\mathbb{P}\left[\sum_{\alpha \in A_j^2} \mathit{Boost}_{i,\alpha} \geq (\mu_{\text{other}} + \Delta_\mu/3)\, m\right] = \mathbb{P}\left[\sum_{\alpha \in A_j^2} \mathit{Boost}_{i,\alpha} \geq \left(1 + \frac{\Delta_\mu}{3\,\mu_{\text{other}}}\right) \mu_{\text{other}}\, m\right] \\
&\leq \exp\left(-\frac{\left(\frac{\Delta_\mu}{3\,\mu_{\text{other}}}\right)^2 \mu_{\text{other}}\, m}{\left(2 + \frac{\Delta_\mu}{3\,\mu_{\text{other}}}\right) c}\right) = \exp\left(-\frac{\Delta_\mu\, m}{\left(\frac{18\,\mu_{\text{other}}}{\Delta_\mu} + 3\right) c}\right) \\
&\leq \exp\left(-\frac{C_{\sigma^2,c}\, r\, m}{\left(\frac{18\,c\,r}{C_{\sigma^2,c}\, r} + 3\right) c}\right) = \exp\left(-\frac{C_{\sigma^2,c}^2}{(18\,c + 3\,C_{\sigma^2,c})\; c}\, r\, m\right).
\end{aligned}$$

By a union bound, the probability that an events $\sum_{\alpha \in A_i^2} Boost_{i,\alpha} \le (\mu_{\text{own}} - \Delta_\mu/3)\, m$ occurs for any i or that an event $\sum_{\alpha \in A_j^2} Boost_{i,\alpha} \ge (\mu_{\text{other}} + \Delta_\mu/3)\, m$ occurs for any $i \neq j$ is at most

$$n \exp\left(-\frac{C_{\sigma^2,c}^2}{18\,c^2}\, r\, m\right) + n\,(n-1) \exp\left(-\frac{C_{\sigma^2,c}^2}{\left(18\,c + 3\,C_{\sigma^2,c}\right)\, c}\, r\, m\right) \le n^2 \exp\left(-\frac{C_{\sigma^2,c}^2}{\left(18\,c + 3\,C_{\sigma^2,c}\right)\, c}\, r\, m\right).$$

When none of these events occur, fixing any pair of agents $i \neq j$,

$$\sum_{\alpha \in A_i^2} Boost_{i,\alpha} - \sum_{\alpha \in A_j^2} Boost_{i,\alpha} \ge (\mu_{\text{own}} - \Delta_\mu/3) m - (\mu_{\text{other}} + \Delta_\mu/3)\, m = \Delta_\mu/3 \cdot m \ge \frac{C_{\sigma^2,c}}{3}\, r\, m.$$

Therefore we have finished the proof where $const\,(\sigma^2, c) = \frac{C_{\sigma^2,c}^2}{\left(18\,c + 3\,C_{\sigma^2,c}\right) c}$, and $K_{\sigma^2,c} := \frac{C_{\sigma^2,c}}{18}$. □

4.2 Bounded Envy of Base Utilities of Suballocation $\{A_i^2\}_{i \in N}$

Next, we bound the inequality between the base utilities for the bundles of the suballocation $\{A_i^2\}_{i \in N}$. Recall that the items in this suballocation are allocated based only on their boosts, independently of their base utilities. As a result, for each agent i, the random variables $\sum_{\alpha \in A_j^2} \underline{u}_i(\alpha)$ are symmetric across all agents j (including $j = i$). Since these random variables have equal expected values, differences between the random variables are limited to random deviations which can be bounded with high probability. We defer the proof of the lemma to Appendix C.3.

Lemma 5. *With probability at least* $1 - n^2 \exp\left(-const\,(\sigma^2, c)\, r\, m\right)$*, it holds for all agents* $i \neq j$ *that*

$$\underline{u}_i(A_i^2) - \underline{u}_i(A_j^2) \ge -4\, K_{\sigma^2,c}\, r\, m,$$

where $K_{\sigma^2,c}$ *was defined in Lemma 4.*

4.3 Putting the Lemmas Together for the Main Theorem

To prove the main theorem, it remains to combine the above bounds on the envy in the suballocation $\{A_i^2\}_{i \in N}$ with the fact that the suballocation $\{A_i^1\}_{i \in N}$ satisfies EF1 by construction, and to show that, in the two scenarios described in the theorem statement, the probability of one of the concentration events not holding and possible positive envy both vanish.

Proof of Theorem 3. Using a union bound, the events described in Lemmas 4 and 5 occur simultaneously with probability at least

$$\begin{aligned} & 1 - n^2 \exp\left(-const\,(\sigma^2, c)\, r\, m\right) - n^2 \exp\left(-const\,(\sigma^2, c)\, r\, m\right) \ge 1 - 2\, n^2 \exp\left(-const\,(\sigma^2, c)\, r\, m\right) \\ & \ge 1 - \exp\left(-const\,(\sigma^2, c)\, r\, m + 2\, \ln n + \ln 2\right) \\ & \ge 1 - \exp\Big(-\underbrace{const\,(\sigma^2, c) \cdot (1 - 1/e)}_{\text{denote this by } B_{\sigma^2,c}} \cdot \min(p, 1/n) \cdot m + 2\, \ln n + \ln 2\Big) \qquad (2) \\ & = 1 - \exp\left(-B_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m + 2\, \ln n + \ln 2\right), \qquad (3) \end{aligned}$$

where the last inequality follows from our lower bound on r (Lemma 11). Assuming that these events indeed occur, for any $i \neq j$

$$u_i(A_j) - u_i(A_i) = \left(\underline{u}_i(A_j^1) + \underline{u}_i(A_j^2) + \sum_{\alpha \in A_j^2} Boost_{i,\alpha}\right) - \left(\underline{u}_i(A_i^1) + \underline{u}_i(A_i^2) + \sum_{\alpha \in A_i^2} Boost_{i,\alpha}\right)$$

$$= \underbrace{\left(\underline{u}_i(A_j^1) - \underline{u}_i(A_i^1)\right)}_{\le 1,\text{ by EF1}} + \underbrace{\left(\underline{u}_i(A_j^2) - \underline{u}_i(A_i^2)\right)}_{\le 4\,K_{\sigma^2,c}\,r\,m\text{ by Lemma 5}} + \underbrace{\left(\sum_{\alpha\in A_j^2} Boost_{i,\alpha} - \sum_{\alpha\in A_i^2} Boost_{i,\alpha}\right)}_{\le -6\,K_{\sigma^2,c}\,r\,m,\text{ by Lemma 4}}$$
$$\le 1 - 2\,K_{\sigma^2,c}\,r\,m \le 1 - K_{\sigma^2,c}\cdot\min(p, 1/n)\cdot m, \tag{4}$$

where the last inequality follows from our lower bound on r and the fact that $1 - 1/e \approx 0.63 \ge 1/2$.

To conclude the proof, it will suffice to show that, in both scenarios ($n \to \infty$ and bounded n) from the theorem statement,

$$B_{\sigma^2,c}\cdot\min(p, 1/n)\cdot m - 2\,\ln n - \ln 2 \to \infty, \tag{5}$$

which by Eq. (3) implies that the concentration events in the lemmas occur with high probability, and to show that

$$\min(p, 1/n)\cdot m \to \infty, \tag{6}$$

which by Eq. (4) implies that, when these concentration events occur, envy vanishes.

In the first scenario, i.e., if $n \to \infty$, $m \to \infty$, $m \ge 3/B_{\sigma^2,c}\,n\ln n$, and $p \ge \frac{3/B_{\sigma^2,c}\,\ln n}{m}$, then $\min(p, 1/n) \ge \min\left(\frac{3/B_{\sigma^2,c}\,\ln n}{m}, \frac{3/B_{\sigma^2,c}\,\ln n}{m}\right) = \frac{3/B_{\sigma^2,c}\,\ln n}{m}$, which shows that the terms in Eqs. (5) and (6) both go to infinity scaling in $\Omega(\ln n)$. This proves the claim for the first scenario, setting $const\,(\sigma^2, c) := 3/B_{\sigma^2,c}$.

In the second scenario, i.e., if $m \to \infty$, n is bounded above, and $p \in \omega(1/m)$, then $\min(p, 1/n)\cdot m \in \min\left(\omega(1/m), \Theta(1)\right)\,m \to \infty$ (satisfying Eq. (6)), and term still goes to infinity when multiplied by a positive constant and when subtracting $O(1)$ terms from it (as in Eq. (5)). This proves the claim for the second scenario and concludes the proof. □

5 NONEXISTENCE OF FAIR ALLOCATIONS AND TIGHTNESS

Our existence theorem (Theorem 3) gave us *positive results* on the boost probability required for envy-freeness, in the sense that whenever the boost probability p lies above the given bounds, envy-free allocations exist with high probability. In this section, we give complementary *negative results* that show that this analysis is tight, i.e., if p lies below the bounds in Theorem 3 by more than a constant factor, some sequence of instances shows that envy-free allocations need no longer exist with high probability.

Recall that our theorem required the probability of boosts to be at least $const\,(\sigma^2, c)\;\ln n/m$ when $n \to \infty$ or to be in $\omega(1/m)$ if n is bounded. Here, we will show that, for any boost distribution $\mathcal{D}$,

- if $m \to \infty$, $n \to \infty$, and $p \le 1/3\;\ln n/m$, there is a sequence of base utilities such that, with high probability, no envy-free allocation exists (Proposition 6), and
- if $m \to \infty$ and $p \in O(1/m)$, there is a sequence of base utilities on which, with $\Omega(1)$ probability, no envy-free allocation exists (Proposition 7).

Note that these negative results are very broad, in the sense that they are tight with our positive results not just in a few instances, but across all boost distributions $\mathcal{D}$ and a wide range of joint evolutions of m, n, and p.

Our nonexistence results will all be based on the archetype of allocation instances without envy-free allocations: instances in which (at least before the boosts) only a single item is valuable.

Proposition 6. *Fix any boost distribution $\mathcal{D}$ and any infinite sequence of triples (m, n, p) such that $m \to \infty$, $n \to \infty$, and $p \le 1/3\,\frac{\ln n}{m}$. Then, there exists an infinite sequence of instances with these values of m, n, and p such that, with high probability, no envy-free (and even no proportional) allocation exists.*

Proof. Whenever $m < n$, we can select any instance in which all agents have positive base value for all items. Then, envy-free allocation cannot exist at all since any allocation must leave some agent without items, who then envies any agent who receives a positive number of items.

Thus, we may assume that $m \geq n$. For these triples (m, n, p), we construct an instance with one special item α_*, for which each agent has a base value of 1, and with $(m-1)$ items for which each agent has base value 0. Note that, by $m \geq n$, it holds that $p \leq 1/3\,\frac{\ln n}{m} \leq 1/3\,\frac{\ln m}{m} \to 0$. In particular, it holds for all large enough m that $p \leq 0.79$, which means that $1 - p \geq e^{-2p}$, which we assume from here on.

For each agent, the probability that they have no boosts for any of the items is $(1-p)^m \geq e^{-2pm} \geq e^{-2/3 \ln n} = \frac{1}{n^{2/3}}$. Since this event is independent between agents, and the expected number of agents without any boosts is $n \cdot \frac{1}{n^{2/3}} = n^{1/3} \to \infty$, concentration of measure implies that, with high probability, there are at least two agents without boosts. In this case, in any allocation of the items, there is an agent who has no boosts and does not receive α_*. This agent necessarily envies the recipient of α_* (and furthermore does not receive their proportional share), which shows that, with high probability, there indeed exists no envy-free (or even just proportional) allocation. □

In Appendix D, we prove another negative result, based on the same idea but without the assumption that n goes to infinity, which is tight with our positive result for bounded n.

Proposition 7. *Fix any boost distribution $\mathcal{D}$ and any infinite sequence of triples (m, n, p) such that $m \to \infty$ and $p \in O(1/m)$. Then, there exists an infinite sequence of instances with these values of m, n, and p such that, with $\Omega(1)$ probability, no envy-free (and even no proportional) allocation exists.*

The two propositions above show that the bounds in Theorem 3 are the most permissive bounds on p one could hope for, up to constants, in both scenarios "$m \to \infty$, $n \to \infty$" and "$m \to \infty$, $n \in O(1)$". One way to interpret the transition from non-existence to existence is that, when p is large enough such that, with high probability, every agent receives a boost, then increasing the boost probability by a constant factor is sufficient to ensure the existence of an envy-free allocation with high probability. One way in which a slightly stronger statement than Theorem 3 might be true is by relaxing the constraint "$m \geq const(\sigma^2, c)\, n \log n$" in the first case. By the nonexistence result of Manurangsi and Suksompong [33] for the distributional model, we know that this constraint on m can be at most relaxed by a factor of $O(\ln \ln n)$, even for $p = 1$, but we do not know if envy-free allocations exist within this log-log gap. It seems that any positive results in this space would require a different algorithmic approach since the deviations of the base utilities for $\{A_i^2\}_i$ can no longer be bounded in $O(r\, m)$ as we do in Lemma 5, and might therefore outweigh the negative envy produced by the boosts.

6 EFFICIENTLY FINDING ENVY-FREE ALLOCATIONS

While the existence proof for Theorem 3 is constructive in a mathematical sense, it does not describe a proper allocation algorithm. What prevents the procedure in the proof from being executed as an algorithm is that the procedure allocates items based on their base utility and their boosts, whereas an allocation algorithm can only observe the utilities, not how these utilities decompose into base utilities and boosts. Since finding envy-free allocations is NP-hard in the worst case (by a straightforward reduction from Partition for $n = 2$), it is not obvious that the envy-free allocations guaranteed to exist in the smoothed model can also be found efficiently.

In this section, we show that envy-free allocations can indeed be found in polynomial time, in (up to constants) the full parameter range allowed by Theorem 3. The key idea is to find a fractional partial allocation that can play the role of $\{A_i^2\}_{i \in N}$, i.e., that only allocates $O(r\, m)$ items to each agent and has $-\Omega(r\, m)$-bounded envy, where we recall that $r = (1 - (1-p)^n)/n$. Suppose, for now, that the algorithm "guesses" a value ρ that gets close to the unknown parameter r in the sense that $\rho \in [r, 2\,r)$. Then, a substitute allocation for $\{A_i^2\}_{i \in N}$ can be found by the following linear program, where the variables $x_{i,\alpha}$ indicate the fraction of item α allocated to agent i:

$$
\begin{aligned}
&\text{maximize } t \\
&\text{s.t.} \sum_{\alpha \in M} x_{i,\alpha} \leq 2\,\rho\,m && \forall i \in N \quad (7)\\
&\sum_{\alpha \in M} u_i(\alpha)\,x_{i,\alpha} \geq t + \sum_{\alpha \in M} u_i(\alpha)\,x_{j,\alpha} && \forall i \neq j \quad (8)\\
&\sum_{i \in N} x_{i,\alpha} \leq 1 && \forall \alpha \in M \quad (9)\\
&x_{i,\alpha} \geq 0 && \forall i \in N, \alpha \in M. \quad (10)
\end{aligned}
$$

As we sketch here, our existing results from Section 4 imply the existence of solutions with high objective value t: Indeed, if we were to set $x_{i,\alpha} := \mathbb{1}\{\alpha \in A_i^2\}$ for all i and α and $t := 2\,K_{\sigma^2,c}\,r\,m$, Eqs. (9) and (10) would directly hold, and these values would also satisfy Eq. (8), by Lemmas 4 and 5, assuming that certain tail events do not occur. As a simple concentration bound in our proof will show, $\sum_{\alpha \in M} x_{i,\alpha} = |A_i^2|$ is at most $2\,\rho\,m$ with high probability, satisfying Eq. (7). Thus, with high probability, optimizing the LP will return a solution $\{x_{i,\alpha}^*\}, t^*$ with objective value $t^* \geq 2\,K_{\sigma^2,c}\,r\,m$.

We then randomly round the fractional allocation, by allocating each item α to each agent i with probability $x_{i,\alpha}^*$, independently across items.[7] We will show that, due to concentration, the resulting partial (non-fractional) allocation still has -1-bounded envy with high probability. This step crucially depends on the fact that the expected number of items allocated to each agent in this partial allocation (smaller than $2\,\rho\,m$ by Eq. (7)) is not too large relative to the envy gap t^*; if ρ was guessed too large, the variance of the random allocation could outweigh this envy gap and lead to envy that is not -1-bounded. After the random-allocation phase, the algorithm uses a polynomial-time subroutine such as round robin to allocate the remaining items in an EF1 way. If the randomly rounded suballocation indeed had -1-bounded envy, combining both partial allocations will be envy-free.

To specify the algorithm, it remains to describe how the algorithm "guesses" an appropriate value of ρ. To do so, it simply runs the above steps for logarithmically many values $\rho = 1/m, 2/m, 4/m, 8/m, \dots, 1$. Each of these runs produces an allocation, for which the algorithm computes in polynomial time the maximum envy between any pair of agents. The algorithm then returns the allocation for which this maximum envy is lowest. As we show in Appendix E, the returned allocation is likely to be envy-free:

Theorem 8. *The randomized algorithm described above runs in polynomial time, and,*

- *if $n \to \infty$ and $m \to \infty$, $m \geq const(\sigma^2, c)\, n \ln n$, and $p \geq \frac{const(\sigma^2, c) \ln n}{m}$, or*
- *if $m \to \infty$, n is bounded, and $p \in \omega(1/m)$,*

it returns an envy-free allocation with high probability.

Let us clarify to which events the probability in Theorem 8 refers to. With high probability, the boosts are such that not only an envy-free allocation exists, but also the resulting allocation instance is such that the randomized algorithm would with high probability, over its own randomness, output an envy-free allocation. Of course, given such an allocation instance, the success probability of the algorithm can be made arbitrarily close to 1, by repeatedly running it with fresh independent randomness.

[7] If $\sum_{i \in N} x_{i,\alpha}^* < 1$ for an item α, α will remain unallocated in this phase with positive probability.

7 EFFICIENTLY FINDING PARETO-OPTIMAL AND PROPORTIONAL ALLOCATIONS

In this section, we give a second example for a hard computational task that becomes easier in the smoothed model: finding allocations that are simultaneously proportional and Pareto-optimal (from here on: *PROP+PO*). Since any envy-free allocation is also proportional, it follows from the previous sections that allocations satisfying just proportionality exist (with high probability), and that we can find them efficiently; and this analysis is tight since our negative results in Propositions 1, 6, and 7 also rule out proportionality. Moreover, the existence of a proportional allocation $\mathcal{A}$ implies that there is a PROP+PO allocation: Indeed, if $\mathcal{A}$ is not Pareto-optimal itself, one verifies that any $\mathcal{A}'$ that Pareto dominates $\mathcal{A}$ must also be proportional. Without loss of generality, we may choose such an $\mathcal{A}'$ to be maximal in the sense that $\mathcal{A}'$ is not itself Pareto dominated, which implies that $\mathcal{A}'$ is Pareto-optimal and PROP+PO. From a computational point of view, however, finding such PROP+PO allocations requires superpolynomial time, even for an algorithm that already receives a proportional allocation as part of its input, unless P = NP (Appendix F).

For our smoothed model, by contrast, we give a polynomial-time algorithm that, with high probability, finds PROP+PO allocations in the scenarios described in Theorem 3. Our algorithm finds a Pareto-optimal *fractional* allocation that exceeds proportionality, in the sense that each agent's utility for their own bundle exceeds their proportional share by at least an amount t. Then, our algorithm rounds this fractional allocation into a non-fractional allocation that is still Pareto-optimal and exceeds proportionality by not much less than t; for large enough t, this implies PROP+PO. Reusing the random rounding procedure from Section 6 would give unsatisfying bounds for this application, since this fractional allocation might give some agents so many items that the variance of the random allocation would outweigh the gap t. Fortunately, the target axioms PROP+PO allow us to use a deterministic rounding scheme, in which the gap t only has to be a constant for the rounded allocation to be proportional. As a result, the algorithm can provide PROP+PO under the assumptions of Theorem 3:

Theorem 9. *There is a (deterministic) polynomial-time allocation algorithm whose returned allocation $\{A_i\}_{i \in N}$ is guaranteed to be Pareto-optimal and such that, in both scenarios of Theorem 3, the allocation is proportional with high probability over the random boosts.*

Proof. For a fractional allocation in which each agent i receives a fraction $x_{i,\alpha} \in [0, 1]$ of each item α, we say that an agent i's *utility* in this fractional allocation is $\sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha}$.

First, the algorithm solves a linear program to determine the largest amount t^* by which all agents' utilities can exceed proportionality in a fractional allocation:

$$
\begin{aligned}
t^* := \text{maximize } & t \\
\text{s.t. } & \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} \geq u_i(M)/n + t && \forall i \in N \\
& \sum_{i \in N} x_{i,\alpha} = 1 && \forall \alpha \in M \\
& x_{i,\alpha} \geq 0 && \forall i \in N, \alpha \in M.
\end{aligned}
$$

Note that this linear program is feasible and bounded. Then, the algorithm solves a second linear program to find the fractional allocation with maximal utilitarian welfare among the fractional allocations exceeding proportionality by t^*:

$$
\begin{aligned}
\text{maximize } & \sum_{i \in N} \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} \\
\text{s.t. } & \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} \geq u_i(M)/n + t^* && \forall i \in N
\end{aligned}
$$

$$\sum_{i \in N} x_{i,\alpha} = 1 \qquad \forall \alpha \in M$$

$$x_{i,\alpha} \geq 0 \qquad \forall i \in N, \alpha \in M.$$

This linear program is feasible due to the choice of t^*, and defines a fractional allocation in which a $x^*_{i,\alpha}$ fraction of each item α is allocated to agent i.

Next, we round this fractional allocation $(x^*_{i,\alpha})_{i,\alpha}$ into a non-fractional allocation $\{A_i\}_{i \in N}$, in such a way that each agent i receives only items α for which $x^*_{i,\alpha} > 0$, and such that each agent i's utility in the non-fractional allocation is at least their utility in the fractional allocation minus the value of a single item. These properties are guaranteed by a well-known deterministic rounding procedure used, for example, by Lenstra et al. [30] and Bezáková and Dani [7], which we sketch below: Consider a bipartite graph, whose left nodes correspond to agents and whose right nodes correspond to items. Let there be an edge between an agent i and an item α iff $x^*_{i,\alpha}$ is *fractional*, i.e., $0 < x^*_{i,\alpha} < 1$. The rounding procedure iteratively modifies the fractional allocation. First, it transfers items along cycles in the bipartite graph until the graph is acyclic, and these steps exactly preserve each agent's utility. Then, it iteratively rounds the remaining fractional $x_{i,\alpha}$, where the acyclicity of the graph allows to ensure that, for each agent i, at most one $x_{i,\alpha}$ gets rounded down. After polynomially many steps, all edges are zero or one, which means that they describe a non-fractional allocation, which our algorithm returns. This procedure satisfies the two properties mentioned at the beginning of the paragraph.

We will now prove that the resulting allocations are Pareto-optimal. Observe that the fractional allocation $(x^*_{i,\alpha})_{i,\alpha}$ is clearly Pareto-optimal (among fractional allocations) since any Pareto dominating fractional allocation would satisfy the constraints of the second linear program and obtain a strictly higher objective. For fractional allocations of goods, it is a classic result that an allocation $(x_{i,\alpha})_{i,\alpha}$ is Pareto-optimal iff it "maximizes a weighted utilitarian welfare" in the sense that there exist positive weights $(\lambda_i)_{i \in N}$ such that, for all i, α with $x_{i,\alpha} > 0$, $i \in \operatorname{argmax}_{j \in N} \lambda_j \cdot u_j(\alpha)$ [35, Lemma 2.3]. Given that our rounding only allocates items α to agents i who originally had $x^*_{i,\alpha} > 0$, the returned allocation is also Pareto-optimal among fractional allocations as certified by the same λ_i, which implies that it is Pareto-optimal among non-fractional allocations.

It remains to show that the allocation $\{A_i\}_{i \in N}$ is likely to be proportional. In either scenario described in its statement, Theorem 3 guarantees that, with high probability, there exists an allocation $\{A_i^{\text{thm}}\}_{i \in N}$ with $\big(1 - K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m\big)$-bounded envy. If this is the case,

$$u_i(A_i^{\text{thm}}) = \frac{\overbrace{u_i(A_i^{\text{thm}}) + \cdots + u_i(A_i^{\text{thm}})}^{n \text{ times}}}{n} \geq \frac{u_i(A_i^{\text{thm}}) + \sum_{j \neq i} u_i(A_j^{\text{thm}}) + (n-1) \cdot \big(K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 1\big)}{n}$$

$$= u_i(M)/n + \tfrac{n-1}{n}\,(K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 1) \geq u_i(M)/n + (K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 1)/2.$$

Thus, t^* is at least $(K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 1)/2$. For any agent i, it must have held for the solution $(x^*_{i,\alpha})_{i,\alpha}$ to the second linear program that $\sum_{\alpha \in M} u_i(\alpha)\, x^*_{i,\alpha} \geq u_i(M)/n + (K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 1)/2$ before rounding. Since, as we showed above, $u_i(A_i)$ is at least $(\sum_{\alpha \in M} u_i(\alpha)\, x^*_{i,\alpha}) - (1 + c)$,

$$u_i(A_i) \geq \left(\sum_{\alpha \in M} u_i(\alpha)\, x^*_{i,\alpha}\right) - (1 + c) \geq u_i(M)/n + (K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m)/2 - (3/2 + c).$$

As we showed in the proof of Theorem 3, $\min(p, 1/n) \cdot m \to \infty$ (Eq. (6)), which means that, for large enough m, $\min(p, 1/n) \cdot m \geq (3 + 2\,c)/K_{\sigma^2,c}$, and then proportionality holds with high probability. □

Compared to the algorithm in Section 6, the allocation algorithm described above is arguably more attractive to run in practice since, even on worst-case instances, Pareto-optimality holds and proportionality cannot be violated by more than the value of a single item (PROP1 [16]) since proportional fractional allocations always exist and hence $t^* \geq 0$. For use in practice, we point out that the objective of the second linear program could be replaced by any other objective maximizing a notion of welfare subject to the constraint; Nash welfare would be a natural choice.

It is an intriguing question for future work when envy-free and Pareto-optimal (*EF+PO*) allocations exist in the smoothed model. In contrast to proportionality, an allocation that Pareto dominates an envy-free allocation need not be envy-free itself, which is why, not just in our smoothed model, positive results for EF+PO tend to be hard to obtain. In particular, it is not clear how one would adapt our PROP+PO proof for such a result, for two reasons: First, it is not clear that there are any fractional allocations that are Pareto-optimal and have $-t$-bounded envy (for large enough t) to begin with. Second, the deterministic rounding approach we used (as well as more refined methods [3, 38]) only bounds each agent's utility for their own bundle — but not for the bundles of others, which would be required to guarantee envy-freeness.

8 DISCUSSION

In this paper, we introduced a smoothed model for additive utilities in a fair-division setting, in which fair allocations are likely to exist and can be found efficiently. Naturally, the take-aways of these results for the real world depend on the degree to which this model reflects utility profiles in practical allocation problems.

Here, we describe one possible interpretation of how our model might capture practical utilities, which we illustrate using the example of splitting an inheritance: Likely, the siblings' values for the items are largely determined by common factors such as the items' cost or usefulness (so agents will largely have similar values for the same item, which tends to make envy-free allocations rarer). There might also be some differences between the agents and items (who has a large apartment? which items fit only in a large apartment?) that lead to systematic patterns between agent–item values, and we will define the worst-case base utilities as the sum of all the above effects. Perhaps, however, there is an additional "idiosyncratic" component to the utilities that gives each sibling a bit of added nostalgic value for a few items based on their individual experiences, in ways that are not predictable from the general patterns between the agents' valuations.

If agents have such idiosyncratic value components, then our model applies and predicts that envy-free allocations are likely to exist. (Interestingly, the same does not hold for idiosyncratic effects causing dislike rather than fondness, see Appendix A.) Idiosyncratic value gives one possible explanation for why envy-free allocations might frequently exist in practical allocation problems.

In future work, it would be informative to empirically test whether our smoothed model accurately describes the properties of real-world utility profiles that possess envy-free allocations. For example, one could check whether there are partial allocations of only few items with large negative envy, as our model would predict.

ACKNOWLEDGMENTS

We thank Nisarg Shah for analyzing how frequently envy-free allocations exist in allocation problems submitted to Spliddit, and we are grateful to Bailey Flanigan, Dominik Peters, and the anonymous reviewers for insightful discussions and helpful suggestions. This work was partially supported by the National Science Foundation under grants IIS-2147187, CCF-2007080, IIS-2024287, and CCF-1733556; and by the Office of Naval Research under grant N00014-20-1-2488.

REFERENCES

[1] G. Amanatidis, E. Markakis, A. Nikzad, and A. Saberi. 2017. Approximation Algorithms for Computing Maximin Share Allocations. *ACM Transactions on Algorithms (TALG)* 13, 4 (2017), 1–28. https://doi.org/10.1145/3147173

[2] C. Arnsperger. 1994. Envy-Freeness and Distributive Justice. *Journal of Economic Surveys* 8, 2 (1994), 155–186. https://doi.org/10.1111/j.1467-6419.1994.tb00098.x

[3] H. Aziz. 2020. Simultaneously Achieving Ex-Ante and Ex-Post Fairness. In *Proceedings of the 16th International Conference on Web and Internet Economics (WINE)*. 341–355. https://doi.org/10.1007/978-3-030-64946-3_24

[4] H. Aziz and S. Rey. 2020. Almost Group Envy-Free Allocation of Indivisible Goods and Chores. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*. 39–45. https://doi.org/10.24963/ijcai.2020/6

[5] Y. Bai and P. Gölz. 2022. Envy-Free and Pareto-Optimal Allocations for Agents with Asymmetric Random Valuations. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. https://arxiv.org/pdf/2109.08971.pdf.

[6] S. Barman, S. K. Krishnamurthy, and R. Vaish. 2018. Finding Fair and Efficient Allocations. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*. 557–574. https://doi.org/10.1145/3219166.3219176

[7] I. Bezáková and V. Dani. 2005. Allocating Indivisible Goods. *ACM SIGecom Exchanges* 5, 3 (2005), 11–18. https://doi.org/10.1145/1120680.1120683 Version with proofs at http://people.cs.uchicago.edu/~ivona/PAPERS/maxmin.pdf.

[8] A. Blum, I. Caragiannis, N. Haghtalab, A. D. Procaccia, E. B. Procaccia, and R. Vaish. 2017. Opting into Optimal Matchings. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2351–2363. https://doi.org/10.1137/1.9781611974782.155

[9] A. Blum and P. Gölz. 2021. Incentive-Compatible Kidney Exchange in a Slightly Semi-Random Model. In *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*. https://doi.org/10.1145/3465456.3467622

[10] A. Blum and J. Spencer. 1995. Coloring Random and Semi-Random k-Colorable Graphs. *Journal of Algorithms* 19, 2 (1995), 204–234. https://doi.org/10.1006/jagm.1995.1034

[11] A. Bogomolnaia, H. Moulin, F. Sandomirskiy, and E. Yanovskaya. 2017. Competitive Division of a Mixed Manna. *Econometrica* 85, 6 (2017), 1847–1871. https://doi.org/10.3982/ECTA14564

[12] E. Budish. 2011. The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103. https://doi.org/10.1086/664613

[13] I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang. 2019. The Unreasonable Fairness of Maximum Nash Welfare. *ACM Transactions on Economics and Computation* 7, 3 (2019), 1–32. https://doi.org/10.1145/3355902

[14] B. R. Chaudhury, J. Garg, and K. Mehlhorn. 2020. EFX Exists for Three Agents. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*. 1–19. https://doi.org/10.1145/3391403.3399511

[15] G. Chèze. 2021. Envy-Free Cake Cutting: A Polynomial Number of Queries with High Probability. arXiv:2005.01982

[16] V. Conitzer, R. Freeman, and N. Shah. 2017. Fair Public Decision Making. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*. 629–646. https://doi.org/10.1145/3033274.3085125

[17] J. P. Dickerson, J. Goldman, J. Karp, A. D. Procaccia, and T. Sandholm. 2014. The Computational Rise and Fall of Fairness. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*. 1405–1411. https://dl.acm.org/doi/10.5555/2893873.2894091

[18] A. Farhadi, M. Ghodsi, M. T. Hajiaghayi, S. Lahaie, D. Pennock, M. Seddighin, S. Seddighin, and H. Yami. 2019. Fair Allocation of Indivisible Goods to Asymmetric Agents. *Journal of Artificial Intelligence Research* 64 (2019), 1–20. https://doi.org/10.1613/jair.1.11291

[19] U. Feige. 2021. Introduction to Semi-Random Models. In *Beyond the Worst-Case Analysis of Algorithms*, T. Roughgarden (Ed.). Cambridge University Press, 189–211. https://doi.org/10.1017/9781108637435.013

[20] U. Feige and J. Kilian. 1998. Heuristics for Finding Large Independent Sets, with Applications to Coloring Semi-Random Graphs. In *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 674–683. https://doi.org/10.1109/SFCS.1998.743518

[21] M. Fleurbaey. 2021. Economics and Economic Justice. In *The Stanford Encyclopedia of Philosophy* (summer 2021 edition ed.), E. N. Zalta (Ed.). https://plato.stanford.edu/archives/sum2021/entries/economic-justice/

[22] R. Freeman, S. Sikdar, R. Vaish, and L. Xia. 2019. Equitable Allocations of Indivisible Goods. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. 280–286. https://doi.org/10.24963/ijcai.2019/40

[23] J. Goldman and A. D. Procaccia. 2014. Spliddit: Unleashing Fair Division Algorithms. *ACM SIGecom Exchanges* 13, 2 (2014), 41–46. https://doi.org/10.1145/2728732.2728738

[24] A. Gupta and S. Singla. 2021. Random-Order Models. In *Beyond the Worst-Case Analysis of Algorithms*, T. Roughgarden (Ed.). Cambridge University Press, 234–258. https://doi.org/10.1017/9781108637435.015

[25] D. K. Herreiner and C. D. Puppe. 2009. Envy Freeness in Experimental Fair Division Problems. *Theory and Decision* 67 (2009), 65–100. https://doi.org/10.1007/s11238-007-9069-8

[26] C. Karande, A. Mehta, and P. Tripathi. 2011. Online Bipartite Matching with Unknown Distributions. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing (STOC)*. 587–596. https://doi.org/10.1145/1993636.1993715

[27] C. Kenyon. 1996. Best-Fit Bin-Packing with Random Order. In *Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 359–364. https://dl.acm.org/doi/10.5555/313852.314083

[28] A. Kolla, K. Makarychev, and Y. Makarychev. 2011. How to Play Unique Games against a Semi-Random Adversary: Study of Semi-Random Models of Unique Games. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*. 443–452. https://doi.org/10.1109/FOCS.2011.78
[29] D. Kurokawa, A. D. Procaccia, and J. Wang. 2016. When Can the Maximin Share Guarantee Be Guaranteed?. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*. 523–529. http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12282
[30] J. K. Lenstra, D. B. Shmoys, and É. Tardos. 1990. Approximation Algorithms for Scheduling Unrelated Parallel Machines. *Mathematical programming* 46, 1 (1990), 259–271. https://doi.org/10.1007/BF01585745
[31] R. J. Lipton, E. Markakis, E. Mossel, and A. Saberi. 2004. On Approximately Fair Allocations of Indivisible Goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce (EC)*. 125–131. https://doi.org/10.1145/988772.988792
[32] P. Manurangsi and W. Suksompong. 2017. Asymptotic Existence of Fair Divisions for Groups. *Mathematical Social Sciences* 89 (Sept. 2017), 100–108. https://doi.org/10.1016/j.mathsocsci.2017.05.006
[33] P. Manurangsi and W. Suksompong. 2020. When Do Envy-Free Allocations Exist? *SIAM Journal on Discrete Mathematics* 34, 3 (2020), 1505–1521. https://doi.org/10.1137/19M1279125
[34] P. Manurangsi and W. Suksompong. 2021. Closing Gaps in Asymptotic Fair Division. *SIAM Journal on Discrete Mathematics* 35, 2 (2021), 668–706. https://doi.org/10.1137/20M1353381
[35] F. Sandomirskiy and E. Segal-Halevi. 2020. Efficient Fair Division with Minimal Sharing. (2020). arXiv:1908.01669
[36] N. Shah. 2021. Personal Communication.
[37] D. A. Spielman and S.-H. Teng. 2004. Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time. *J. ACM* 51, 3 (2004), 385–463. https://doi.org/10.1145/990308.990310
[38] A. Srinivasan. 2008. Budgeted Allocations in the Full-Information Setting. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*. 247–253. https://doi.org/10.1007/978-3-540-85363-3_20
[39] W. Suksompong. 2016. Asymptotic Existence of Proportionally Fair Allocations. *Mathematical Social Sciences* 81 (2016), 62–65. https://doi.org/10.1016/j.mathsocsci.2016.03.007
[40] L. Xia. 2020. The Smoothed Possibility of Social Choice. *Advances in Neural Information Processing Systems (NeurIPS)* 33 (2020), 11044–11055. https://proceedings.neurips.cc/paper/2020/file/7e05d6f828574fbc975a896b25bb011e-Paper.pdf
[41] L. Xia. 2021. The Semi-Random Satisfaction of Voting Axioms. *Advances in Neural Information Processing Systems (NeurIPS)* 34 (2021). https://proceedings.neurips.cc/paper/2021/file/3083202a936b7d0ef8b680d7ae73fa1a-Paper.pdf
[42] D. Zeng and A. Psomas. 2020. Fairness-Efficiency Tradeoffs in Dynamic Fair Division. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*. 911–912. https://arxiv.org/pdf/1907.11672.pdf

APPENDIX

A JUSTIFICATION OF MODELING DECISIONS

The smoothed model presented in this paper is the result of multiple modeling decisions. Alternative decisions would have yielded alternative smoothed models, each with different possibilities and limitations in what it can express about the underlying problem of fair item allocation. Here, we describe three such design decisions and a rationale of why we chose our model over the alternative.

Why are boosts positive? It is natural to wonder why we choose our boosts to be positive rather than also allowing negative perturbations. One reason is that nonnegative boosts ensure that we remain in the setting of allocating *goods* (where items have a nonnegative value to all agents), whereas negative boosts might enter the setting of *mixed manna*, where items might have positive or negative value to different agents (in Appendix G, we discuss how our findings generalize to this setting and boosts that need not be positive).

A second reason to focus on positive boosts is that sparse positive boosts are more useful for guaranteeing envy-free allocations than sparse negative boosts: Indeed, consider an instance with n agents, $m = k\,n - 1$ many items for some $k \in \mathbb{N}_{\geq 1}$, and all agent–item base utilities being 1, and suppose that we are applying negative-only boosts to this instance. Then, in any possible allocation, some agent i must receive strictly less than k items, and one can verify that, unless the sum of i's negative boosts $\sum_{\alpha \in M} |Boost_{i,\alpha}|$ is at least $n-1$, i must envy another agent. If, say, $p = n^{0.99}/m$ as n and m go to infinity, standard concentration bounds show that, with high probability, all agents will have $o(n)$ total boosts, which means that envy-free allocations are unlikely to exist. By contrast, Theorem 3 shows that, for positive boosts, envy-free allocations exist when p is larger than a constant factor times $\ln n/m$, and that therefore a much smaller amount of boosts suffices. An alternative model permitting negative-only boosts would have obscured this existence by the above impossibility.

Though we could consider boosts that are positive and negative (both sufficiently often and with sufficient magnitude), there is an important conceptual principle that "additional perturbations make things no harder." Suppose, for example, that the boost distribution is uniform on [-1, 1], so 50% of the boosts are negative. Our analysis for envy-freeness still goes through without problems if we treat negative boosts the same way as the worst-case base utilities, i.e.: (1) all items without any boosts or with only negative boosts are allocated via round robin, and (2) all items with some positive boosts are allocated to the agent with the largest (positive) boost.[8] The intuition underlying our modeling choice is that positive boosts alone are sufficient to guarantee envy-freeness, and that in settings with mixed-sign boosts as above, the positive boosts do most of the heavy lifting.

Why a parameter p? In a certain sense, the parameter p is superfluous in our model since it could be incorporated into the boost distribution $\mathcal{D}$ by putting a point mass of $1-p$ on 0. However, having p as an explicit parameter is informative, since it allows us to say, when p is small, that a large fraction of agent–item utilities are identical to the worst-case base utilities. Moreover, fixing $\mathcal{D}$ and only varying p allows us to express a family of boost distributions parameterized by a single number, and to express that the expected value and the variance of the $Boost_{i,\alpha}$ may decrease as m increases while still guaranteeing envy-free allocations (which is not surprising, because there will be more items to which the boost is applied). Of course, p is not the only way of parameterizing a family of boost distributions, and only models one interpretation of "decreasing the boosts".

[8]While this simple argument is somewhat limited in how generally it is applicable, the basic strategy of treating additional randomness as part of the worst-case input is more broadly applicable.

Why $m \to \infty$? The general goal of the study of fair (e.g., envy-free) allocations in a smoothed model is to study for various values of n and m and various profiles of base utilities, what the relation is between the extent of random perturbations to the base utilities and the probability of existence of fair envy-free allocations. Within such a framework, it is desirable to identify conditions under which the probability of existence of envy-free allocations tends to 1. We expressed such conditions with a parameterization of $m \to \infty$. An alternative way of trying to express such conditions is as follows: fix a worst-case allocation instance (i.e., numbers n, m, and a profile of base utilities), and then, lower-bound how fast the probability of existence of envy-free allocations converges to 1, as we vary the random perturbations to be increasingly conducive to the existence of such allocations. We chose not to take this approach, for reasons explained below.

A major issue with this approach is that, considering a wide class of boost distributions, there are worst-case allocation instances for which, under any such boost distribution, envy-free allocations only exist with a probability bounded away from one: Indeed, for any m and n, consider the instance where all n agents have base utility 1 for one special item α_* and base utility zero otherwise. Without boosts, clearly no envy-free allocation exists. Now, consider any boost distribution $\mathcal{D}$ supported on $\mathbb{R}_+$ whose probability density function (pdf) is monotone nonincreasing — i.e., boosted utilities closer to the base utility are more likely than utilities further from it, a natural assumption for a random perturbation. With probability $(1/m)^{2\,m} > 0$, agents 1 and 2 will both have a boost above the $(m-1)/m$-th quantile of $\mathcal{D}$ for α_* and a boost below the $1/m$-th quantile for all other items. Then, since the pdf is monotone nonincreasing, the quantile function is convex, and

$$Boost_{i,\alpha_*} \geq quantile_{\mathcal{D}}\big((m-1)/m\big) \geq (m-1)\, quantile_{\mathcal{D}}(1/m) \geq \textstyle\sum_{\alpha \in M \setminus \{\alpha_*\}} Boost_{i,\alpha}$$

for $i = 1$ and $i = 2$. Thus, if agent $i \in \{1, 2\}$ does not receive α_* in a given allocation, i envies the agent who receives α_* even if i receives all other items. Since no allocation can give α_* to both agents, it follows that no allocation is envy-free in this case, which happens with probability at least $m^{-2m} > 0$. Intuitively, such a small probability seems unimportant for all but the smallest numbers of items m, but it means that we cannot use the rate of convergence to 1 of envy-free allocations as a measure, at least not for a single instance, and thus a single m.

We sidestep this issue by considering the limit case of $m \to \infty$, i.e., we study sequences of instances along which the number of items goes to infinity, which is also the approach taken in the literature on distributional models for item allocation. Along such sequences of instances, failure probabilities such as that of m^{-2m} above would vanish, and no longer preclude the existence of envy-free allocations existing with high probability for sufficient amounts of random boosts. Even though this model does not allow us to speak about individual instances anymore, our objective is the same as the motivation of the smoothed-analysis setup: We hope to characterize how much randomness suffices for envy-free allocations, which we mathematically express through how small p may be in relationship to m and n.

B PROOFS OF NEGATIVE RESULTS FOR ROUND ROBIN ALLOCATION

Proposition 1. *For any $\delta > 0$ and $\epsilon \in (0, 1)$, there is a family of instances with two agents, 1 and 2, and $m \to \infty$ items, with a boost probability of $p := m^{-2/3-\delta}$ and the boost distribution Uniform$\big((0, \epsilon)\big)$ such that, with high probability, agent 1 will envy agent 2 in the round robin allocation starting from agent 2.*

Proof. Let the items be labeled $\alpha_1, \ldots, \alpha_m$ (m is odd). For $i = 1, \ldots, m^{2/3}$, set the base utilities as

$$\underline{u}_1(\alpha_i) = 1 - \frac{i-1}{m^{2/3}-1}\,\epsilon \quad \underline{u}_2(\alpha_i) = 1 - \frac{i-1}{m^{2/3}-1}\,(1-\epsilon),$$

and for all $i = m^{2/3} + 1, \ldots, m$ set the base utilities as

$$\underline{u}_1(\alpha_i) = 1 - \epsilon \quad \underline{u}_2(\alpha_i) = 0.$$

We will assume that ties between equal utilities are broken in favor of items with lower index. Now we prove that running round robin on the instances above and giving the first turn to agent 2, agent 1 will envy agent 2 with high probability.

By a union bound, the probability that at least one out of the items $\alpha_1, \ldots, \alpha_{m^{2/3}}$ is boosted for at least one agent is at most $2\, m^{2/3}\, p = 2\, m^{-\delta}$, which goes to zero for large enough instances. Thus, we may assume from here on that none of these items are boosted.

Also, observe that, since $\underline{u}_2(\alpha_i) \geq \epsilon$ for all $i \leq m^{2/3}$, the boosts of other items do not change that, for all $\ell = 1, \ldots, m^{2/3}$, α_ℓ is agent 2's unique item with the ℓth largest utility.

Next, for some $\ell = 2, \ldots, m^{2/3} - 1$, consider the event that more than one item in $\{\alpha_i\}_{i > m^{2/3}}$ has a boosted utility for agent 1 between $\underline{u}_1(\alpha_{\ell-1})$ and $\underline{u}_1(\alpha_{\ell+1})$. For this to happen, two of these items must be boosted and have their uniformly-drawn boost fall in the same subset of probability measure $2/(m^{2/3} - 1)$. A simple union bound over the pairs of items shows that this happens with probability at most $m^2\,(2\,p/(m^{2/3} - 1))^2 = 4\, m^{2/3 - 2\,\delta}/(m^{2/3} - 1)^2$. With high probability, this is not the case for any $\ell = 2, \ldots, m^{2/3} - 1$ by union bound, and we assume this going forward.

Under these assumptions, the round robin process looks as follows from agent 1's perspective: Agent 2 takes α_1, which agent 1 values at 1. From here on, we think of the item picks as being paired up: agent 1 picks an item $\alpha_{X_1^1}$, which we pair with agent 2's subsequent pick $\alpha_{X_1^2}$; then agent 1 picks $\alpha_{X_2^1}$ and agent 2 picks $\alpha_{X_2^2}$; and so forth until the items run out. Since there is an odd number of items, agent 2 has the last pick $\alpha_{X^2_{\lfloor m/2 \rfloor}}$. Agent 1's envy can be written as

$$1 - \sum_{\ell=1}^{\lfloor m/2 \rfloor} \underbrace{u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2})}_{\geq 0}.$$

From some point on, only the unboosted items in $\{\alpha_i\}_{i > m^{2/3}}$ will be left, all of which have utility exactly $1 - \epsilon$ to agent 1, which means that most terms $u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2})$ equal zero. The interesting part is what happens before: Both agents have the same preference order over the items $\alpha_1, \ldots, \alpha_{m^{2/3}}$. But, whereas for agent 2, none of the remaining items are interesting until all $m^{2/3}$ low-index items are gone, agent 1 will interleave picking low-index items $\{\alpha_i\}_{i \leq m^{2/3}}$ and boosted items in $\{\alpha_i\}_{i > m^{2/3}}$. Still, our assumption above guarantees that the utilities of the boosted items are sparsely distributed between the utilities of the low-index items for agent 1, which will allow us to analyze this interleaving.

Up to an affine transformation of utilities, this early phase of the round robin coincides with the process described in Lemma 10 below, with the items boosted for agent 1 playing the role of A_2 and the low-index items playing the role of A_1. If T is the round in which the last item is picked that is either boosted for agent 1 or has an index of at most $m^{2/3}$, the lemma guarantees that none of the unboosted items in $\{\alpha_i\}_{i > m^{2/3}}$ were picked strictly before round T and that

$$\sum_{\ell=1}^{\lfloor m/2 \rfloor} u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2}) \leq \epsilon.$$

As discussed above, all $u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2})$ for $\ell > T$ equal zero, which means that agent 1 envies agent 2 by at least $1 - \epsilon > 0$. □

Lemma 10. *Fix $t \in \mathbb{N}$, and let $A_1 \coloneqq \{-t, -t+1, \ldots, -1\}$. Furthermore, fix a set A_2 of real numbers in $[-t, 0] \setminus \mathbb{Z}$ with the property that, for all $\ell = -t+1, \ldots, -2$, it holds that $|\{a \in A_2 \mid \ell - 1 < a <$*

$\ell + 1\}| \leq 1$. *Now, consider the following process with two agents, which take turns removing a single element from* $A_1 \dot{\cup} A_2$ *such that*

- *agent 1 begins,*
- *in agent 1's* ℓ*th turn, they remove the largest remaining element of* $A_1 \cup A_2$ *and call it* a^1_ℓ*, and*
- *in agent 2's* ℓ*th turn, they remove the largest remaining element of* A_1 *and call it* a^2_ℓ*.*

This process is well-defined until no elements remain (i.e., it cannot be the case that in agent 2's turn, A_1 *is empty but* A_2 *is not). Furthermore, denoting the final round by* T *(and, if round* T *is only half-completed, setting* $a^2_T \coloneqq -t$*), it holds that*

$$\sum_{\ell=1}^{T} a^1_\ell - a^2_\ell \leq t. \tag{11}$$

Proof. By induction on t. If $t = 0$, the claim is trivial, and it is also easy to verify for $t = 1$.

From here on, let $t \geq 2$. If $a \geq -t + 1$ for all $a \in A_2$, the induction hypothesis tells us how the process runs up to the last move, in which the player with the next turn will take the element $-t$. Clearly, this process remains well-defined. If agent 1 is the one to take $-t$, this adds a zero term $(-t) - (-t)$ to the sum $\sum_{\ell=1}^{T} a^1_\ell - a^2_\ell$. Else, if agent 2 takes $-t$, this decreases a^2_T by one relative to the induction hypothesis. Since, the right-hand side of Eq. (11) increases by one, this establishes the induction step.

Now consider the remaining case in which there is an $a \in A_2$ such that $a < -t + 1$. By the conditions on A_2, no other $a' \in A_2$ is less than $-t + 2$, which allows us to apply the induction hypothesis for $t - 2$ and for $A_2 \setminus \{a\}$ taking the role of A_2. Clearly, this induction step describes the beginning of the process for the current t, except that three elements of $A_1 \cup A_2$, i.e., $-t + 1, a, -t$, must subsequently be picked. We distinguish two subcases, depending on which agent was the last to move in the process of the induction hypothesis.

If agent 1 was the last to move in the induction hypothesis, agent 2 will pick $-t + 1$, agent 1 will pick a, and agent 2 picks $-t$. This process is well-defined, and increases the left-hand side of Eq. (11) by an amount of $1 + (a - (-t)) \leq 2$ relative to the induction hypothesis, while the right-hand side increases by 2, concluding this case of the induction step.

Else, i.e., if agent 2 was the last to move in the induction hypothesis, agent 1 will pick $-t + 1$, agent 2 will pick $-t$, and then agent 1 will pick a. This process is also well-defined, and also increases the sum by an amount of $1 + (a - (-t))$, which means that the induction step is again established. This concludes the induction. □

Corollary 2. *Fix any* $\delta > 0$, $\epsilon \in (0, 1)$ *and any function* $n : \mathbb{N}_{\geq 1} \to \mathbb{N}_{\geq 2}$ *such that* $n(m) < m^{1/3}$ *for sufficiently large* m*. Then, there is a family of instances with* $m \to \infty$ *items and* $n = n(m)$ *agents, with a boost probability of* $p \coloneqq m^{-2/3-\delta}$ *and the boost distribution Uniform*$\big((0, \epsilon)\big)$ *such that, with high probability, round robin (starting with agent 2) produces an allocation in which agent 1 envies agent 2.*

Proof. Let $n \coloneqq n(m)$ denote the number of agent, then $n < m^{1/3}$. Keep the base utilities of agents 1 and 2 exactly as in the above proof. Set the base utilities of other agents $i = 3, \ldots, n$ to 0 on the first $m^{2/3}$ items and 1 on the rest of the items. Since for all agents $i = 3, \ldots, n$ and items $j = m^{2/3} + 1, \ldots, m$, $\underline{u}_i(\alpha_j) = 1 > \epsilon$, the boosts of the first $m^{2/3}$ items do not change the fact that these agents will always pick items in the set $\{\alpha_j\}_{j > m^{2/3}}$ until the set is empty. Now we prove that running round robin and giving the first turn to agent 2, agent 1 will envy agent 2 with high probability.

Still, let T denote the round in which the last item is picked that is either boosted for agent 1 or has an index of at most $m^{2/3}$. The simulation process in Lemma 10 indicates that agent 2 has not picked any item from $\{\alpha_i\}_{i>m^{2/3}}$ before Tth round, thus agent 2 must have picked at most $m^{2/3}$ items, suggesting that $T \le m^{2/3}$. Hence by the end of Tth round, there will be at least

$$m - m^{2/3} - T \cdot (n-1) \ge m - m^{2/3} - m^{2/3} \cdot (n-1) > m - m^{2/3} - m^{2/3} \cdot (m^{1/3} - 1) = 0$$

items left in $\{\alpha_i\}_{i>m^{2/3}}$, indicating that none of the items in $\{\alpha_i\}_{i\le m^{2/3}}$ has been picked by agent $3, \cdots, n$ before Tth round. Then the only difference between what happens in this instance and the one in the earlier proof with two agents is that some of agent 1's boosted items may be picked by agents $3, \ldots, n$, but this does not affect the analysis of the envy.

Therefore, we still have

$$\sum_{\ell=1}^{T} u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2}) \le \epsilon.$$

As discussed in earlier proof, all $u_1(\alpha_{X_\ell^1}) - u_1(\alpha_{X_\ell^2})$ for $\ell > T$ equal zero, thus agent 1 envies agent 2 by at least $1 - \epsilon > 0$. □

C PROOFS FOR THEOREM 3: ENVY-FREE ALLOCATIONS EXIST

C.1 Bounding r

Lemma 11. *Recall that $r = \frac{1-(1-p)^n}{n}$. It holds that*

$$(1 - 1/e) \min(p, 1/n) \le r \le \min(p, 1/n).$$

Proof. The upper bound follows directly since $r = \frac{1-(1-p)^n}{n} \le \frac{1}{n}$ and $\frac{1-(1-p)^n}{n} \le \frac{1-(1-p\,n)}{n} = p$.

For the lower bound, we will show that, if $p\,n \ge 1$, then $(1-1/e)\,\frac{1}{n} \le r$, and that else, $(1-1/e)\,p \le r$: Assuming that $p\,n \ge 1$,

$$r = \frac{1-(1-p)^n}{n} \ge \frac{1-e^{-p\,n}}{n} \ge \frac{1-e^{-1}}{n}.$$

Else, assuming that $p\,n < 1$,

$$\frac{r}{p} = \frac{1-(1-p)^n}{p\,n} \ge \frac{1-e^{-p\,n}}{p\,n} \ge \min_{x\le 1} \frac{1-e^{-x}}{x} = 1 - \frac{1}{e},$$

where the last inequality follows from the fact that $\frac{1-e^{-x}}{x}$ is monotone decreasing on $\mathbb{R} \setminus \{0\}$. This implies that $r \ge (1-1/e)\,p$ whenever $p \ne 0$, and one verifies that the inequality also holds when $p = 0$. Since, in the two cases, r is always lower-bounded by one out of $(1-1/e)\,p$ and $(1-1/e)\,\frac{1}{n}$, it must be lower-bounded by the minimum of both terms, as claimed. □

C.2 Bounding $\mathbb{E}[\mathcal{D}]$ and $\mathbb{E}[\mathcal{D}] - \mathbb{E}[\mathcal{D}_{2:2}]$ in Terms of σ^2 and c

Lemma 12. *For a distribution $\mathcal{D}$ in $[0, c]$ with variance σ^2, for any range of length $\frac{\sigma}{2}$ in $[0, c]$, the probability of $\mathcal{D}$ that falls in the range is less than $1 - p_{\sigma^2,c} = 1 - \frac{3\,\sigma^2}{4\,(2\,c^2 + c\,\sigma)}$.*

Proof. For a range of length $\frac{\sigma}{2}$ in $[0, c]$, assume $\mathcal{D}$ falls out of the range with probability p and falls in the range with probability $1-p$. The variance of the distribution consists of two parts, those that from the fraction out of the range and those that from the fraction in the range.

- The fraction that falls out of the range can at most contribute $p\,c^2$ to the variance, since the maximum distance between them and the expectation is c.
- For the fraction in the range, first consider the furthest distance $d_{\max}$ of the expectation away from the range. This can be attained when all $1-p$ fraction are on one side of the range and the other p fraction is furthest away in $[0, c]$. Since the p fraction can be at most $c - \frac{\sigma}{2}$ distance away from the $1-p$ fraction, which results in the expectation to be $p\,(c - \frac{\sigma}{2}) < p\,c$ distance away from the range, and thus $d_{\max} < p\,c$. Hence all the $1-p$ fraction in the range is

less than $p\,c + \frac{\sigma}{2}$ distance away from the expectation, which contribute a variance less than $(1-p)\,(p\,c + \frac{\sigma}{2})^2$.

Adding these two parts up, we have

$$\sigma^2 < p\,c^2 + (1-p)\,(p\,c + \frac{\sigma}{2})^2 \le p\,c^2 + (p\,c + \frac{\sigma}{2})^2$$
$$= p\,c^2 + p^2\,c^2 + p\,c\,\sigma + \frac{\sigma^2}{4} \le 2\,p\,c^2 + p\,c\,\sigma + \frac{\sigma^2}{4}$$

Then

$$\frac{3\,\sigma^2}{4} < p\,(2\,c^2 + c\,\sigma) \quad \Rightarrow \quad p > \frac{3\,\sigma^2}{4\,(2\,c^2 + c\,\sigma)} = p_{\sigma^2,c}$$

Therefore, the maximum probability of $\mathcal{D}$ in the range is less than $1 - p_{\sigma^2,c}$. □

Lemma 13. *For a distribution $\mathcal{D}$ in $[0,c]$ with variance σ^2, we have that $\mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}] \ge \sigma\, p^2_{\sigma^2,c}/8$ and $\mathbb{E}[\mathcal{D}] \ge \sigma\, p_{\sigma^2,c}/2$.*

Proof. Let $F_{\mathcal{D}}(x)$ denote the c.d.f. of $\mathcal{D}$, then the c.d.f. of $\mathcal{D}_{2:2}$ is $F^2_{\mathcal{D}}(x)$. It can be deduced that

$$\mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}] = \int_0^c (1 - F^2_{\mathcal{D}}(x))\,dx - \int_0^c (1 - F_{\mathcal{D}}(x))\,dx$$
$$= \int_0^c F_{\mathcal{D}}(x)\,(1 - F_{\mathcal{D}}(x))\,dx$$

Assume $F_{\mathcal{D}}(x_0) = p_{\sigma^2,c}/2$, then by Lemma 12 we have $F_{\mathcal{D}}\,(x_0 + \frac{\sigma}{2}) \le 1 - p_{\sigma^2,c}/2$ (this also implies that $x_0 + \frac{\sigma}{2} < c$). Then in range $[x_0, x_0 + \frac{\sigma}{2}]$, $F_{\mathcal{D}}(x) \ge p_{\sigma^2,c}/2$ and $1 - F_{\mathcal{D}}(x) \ge p_{\sigma^2,c}/2$, hence

$$\mathbb{E}[\mathcal{D}_{2:2}] - \mathbb{E}[\mathcal{D}] = \int_0^c F_{\mathcal{D}}(x)\,(1 - F_{\mathcal{D}}(x))\,dx \ge \frac{\sigma}{2} \cdot (\frac{p_{\sigma^2,c}}{2})^2$$

From Lemma 12 we know that the probability outside $[0, \frac{\sigma}{2}]$ is at least $p_{\sigma^2,c}$, then $\mathbb{E}[\mathcal{D}] \ge \frac{\sigma}{2} \cdot p_{\sigma^2,c}$. □

C.3 Deferred Proof of Lemma 5

Lemma 5. *With probability at least $1 - n^2\, \exp\left(-const(\sigma^2, c)\, r\, m\right)$, it holds for all agents $i \ne j$ that*

$$\underline{u}_i(A_i^2) - \underline{u}_i(A_j^2) \ge -4\, K_{\sigma^2,c}\, r\, m,$$

where $K_{\sigma^2,c}$ was defined in Lemma 4.

Proof. Fix any pair of agents $i \ne j$. Recall that each item α independently is part of A_j^2 with probability r, which means that $\mathbb{E}[\underline{u}_i(A_j^2)] = \sum_{\alpha \in M} r\, \underline{u}_i(\alpha) = r\, \underline{u}_i(M)$. Furthermore, recall that $0 \le \underline{u}_i(\alpha) \le 1$. These observations allow us to apply a Chernoff bound as follows:

$$\mathbb{P}\left[\underline{u}_i(A_j^2) \ge r\,\underline{u}_i(M) + 2\,K_{\sigma^2,c}\,r\,m\right] = \mathbb{P}\left[\underline{u}_i(A_j^2) \ge \left(1 + \frac{2\,K_{\sigma^2,c}\,m}{\underline{u}_i(M)}\right) r\,\underline{u}_i(M)\right]$$
$$\le \exp\left(-\frac{\left(\frac{2\,K_{\sigma^2,c}\,m}{\underline{u}_i(M)}\right)^2 r\,\underline{u}_i(M)}{2 + \frac{2\,K_{\sigma^2,c}\,m}{\underline{u}_i(M)}}\right) = \exp\left(-\frac{2\,K^2_{\sigma^2,c}\,r\,m}{\frac{\underline{u}_i(M)}{m} + K_{\sigma^2,c}}\right) \le \exp\left(-\frac{2\,K^2_{\sigma^2,c}}{1 + K_{\sigma^2,c}}\,r\,m\right).$$

By the same reasoning,

$$\mathbb{P}\left[\underline{u}_i(A_i^2) \le r\,\underline{u}_i(M) - 2\,K_{\sigma^2,c}\,r\,m\right] \le \exp\left(-2\,K^2_{\sigma^2,c}\,r\,m\right).$$

A union bound shows that, with probability at least $1 - n^2 \exp\left(-\frac{2K^2_{\sigma^2,c}}{1+K_{\sigma^2,c}} r\, m\right)$, both $\underline{u}_i(A^2_j), \underline{u}_i(A^2_i)$ are bounded within $2\, K_{\sigma^2,c}\, r\, m$ distance from $r\, \underline{u}_i(M)$ for all $i \neq j$, which implies for all $i \neq j$,

$$\underline{u}_i(A^2_i) - \underline{u}_i(A^2_j) \geq -4\, K_{\sigma^2,c}\, r\, m. \qquad \square$$

D PROOFS FOR NONEXISTENCE

Proposition 7. *Fix any boost distribution $\mathcal{D}$ and any infinite sequence of triples (m, n, p) such that $m \to \infty$ and $p \in O(1/m)$. Then, there exists an infinite sequence of instances with these values of m, n, and p such that, with $\Omega(1)$ probability, no envy-free (and even no proportional) allocation exists.*

Proof. Let $C > 0$ be a constant such that $p \leq C/m$ for sufficiently large m. Because $p \in O(1/m) \to 0$, it holds (for all large enough m) that $1 - p \geq e^{-2p}$. We will use the same base utilities as in the proof of Proposition 6, with one special item α_* having base utility 1 for all agents and with all other items having base utility 0 for all agents. The probability that a given agent has no boosts is

$$(1-p)^m \geq e^{-2\,p\,m} \geq e^{-2\,C},$$

where the inequalities hold for large enough m. When the inequalities hold, the probability that agents 1 and 2 both have no boosts is lower-bounded by the constant $(e^{-2\,C})^2 = e^{-4\,C} > 0$. As in the proof of Proposition 6, there are no envy-free allocations whenever two agents have no boosts, which proves that with probability lower bounded by a function of C, no envy-free allocations exist. The same even holds for proportional allocations. $\square$

E PROOF OF Theorem 8: EFFICIENTLY FINDING ENVY-FREE ALLOCATIONS

Theorem 8. *The randomized algorithm described above runs in polynomial time, and,*

- *if $n \to \infty$ and $m \to \infty$, $m \geq const(\sigma^2, c)\, n \ln n$, and $p \geq \frac{const(\sigma^2,c)\ \ln n}{m}$, or*
- *if $m \to \infty$, n is bounded, and $p \in \omega(1/m)$,*

it returns an envy-free allocation with high probability.

Proof. In Lemma 11, we show that $r \geq (1 - 1/e) \min(p, 1/n)$. Hence, in both scenarios, it holds that $r > 1/m$ when m is large enough. Then during the search for $\rho = 1/m, 2/m, 4/m, \ldots, 1$, there will be one ρ such that $\rho \in [r, 2\,r)$. Consider the allocation that our algorithm returns under this ρ. From Lemmas 4 and 5, we know that with probability at least $1 - 2\,n^2 \exp\left(-\text{const}(\sigma^2, c)\, r\, m\right)$ the suballocation $\{A^2_i\}_{i \in N}$ satisfies:

$$u_i(A^2_i) - u_i(A^2_j) \geq 2\, K_{\sigma^2,c}\, r\, m \quad \forall i \neq j.$$

We can apply Chernoff bound to bound the size of A^2_i for any agent i:

$$\mathbb{P}\left[\left|A^2_i\right| \geq 2\, r\, m\right] \leq \exp\left(-\frac{1}{3}\, r\, m\right).$$

By a union bound, the probability that $|A^2_i| \leq 2\, r\, m \leq 2\, \rho\, m$ for all agent $i \in N$ is at least

$$1 - n \cdot \mathbb{P}\left[\left|A^2_i\right| \geq 2\, r\, m\right] \geq 1 - n \exp\left(-\text{const}(\sigma^2, c)\, r\, m\right).$$

Then with probability at least

$$1 - 2\,n^2 \exp\left(-\text{const}(\sigma^2, c)\, r\, m\right) - n \exp\left(-\text{const}(\sigma^2, c)\, r\, m\right) \geq 1 - 3\,n^2 \exp\left(-\text{const}(\sigma^2, c)\, r\, m\right),$$

both events above happen which, in turn, guarantee the solution of the subroutine linear program: $t^* \geq 2\, K_{\sigma^2,c}\, r\, m > K_{\sigma^2,c}\, r\, m$. Let $\{A^3_i\}_{i \in N}$ denote the integral suballocation after rounding the fractional suballocation. Observe that for any i, j, $u_i(A_j^3) = \sum_{\alpha \in M} u_i(\alpha) X_{j,\alpha}$ where $X_{j,\alpha}$ is a random variable that takes value 1 with probability $x_{j,\alpha}$ and 0 otherwise[9], thus we have

$$\mathbb{E}\left[u_i(A_j^3)\right] = \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha} \le (1+c) \sum_{\alpha \in M} x_{j,\alpha} \le 2\,(1+c)\,\rho\, m,$$

then we can apply Chernoff bound to bound the random deviations of utilities in rounding step for any i, j:

$$\mathbb{P}\left[\left|u_i(A_j^3) - \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha}\right| \ge \frac{1}{3} K_{\sigma^2,c}\, r\, m\right] = \mathbb{P}\left[\left|\sum_{\alpha \in M} \frac{u_i(\alpha)}{1+c} X_{j,\alpha} - \frac{1}{1+c}\sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha}\right| \ge \frac{1}{3\,(1+c)} K_{\sigma^2,c}\, r\, m\right]$$

$$\le 2\,\exp\left(-\frac{1}{3}\cdot\frac{\left(\frac{1}{3\,(1+c)} K_{\sigma^2,c}\, r\, m\right)^2}{2\,\rho\, m}\right) \le 2\,\exp\left(-\frac{1}{3}\cdot\frac{\left(\frac{1}{3\,(1+c)} K_{\sigma^2,c}\, r\, m\right)^2}{4\, r\, m}\right) = 2\,\exp\left(-\frac{K_{\sigma^2,c}^2}{108\,(1+c)^2}\, r\, m\right).$$

Taking a union bound over all pairs $i \ne j$ then

$$\left|u_i(A_j^3) - \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha}\right| \le \frac{1}{3} K_{\sigma^2,c}\, r\, m,\ \forall i \ne j$$
$$\text{with probability} \ge 1 - 2\,n^2\,\exp\left(-\text{const}(\sigma^2, c)\, r\, m\right).$$

If this holds, we have $\forall i \ne j$:

$$u_i(A_i^3) - u_i(A_j^3) \ge u_i(A_i^3) - \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} + \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} - \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha} + \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha} - u_i(A_j^3)$$

$$\ge \underbrace{\sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha} - \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha}}_{\ge t^* > K_{\sigma^2,c}\, r\, m} - \underbrace{\left|u_i(A_i^3) - \sum_{\alpha \in M} u_i(\alpha)\, x_{i,\alpha}\right|}_{\le 1/3\, K_{\sigma^2,c}\, r\, m} - \underbrace{\left|u_i(A_j^3) - \sum_{\alpha \in M} u_i(\alpha)\, x_{j,\alpha}\right|}_{\le 1/3\, K_{\sigma^2,c}\, r\, m} \ge \frac{1}{3} K_{\sigma^2,c}\, r\, m.$$

This happens when the linear program has a solution $t^* \ge 2\,K_{\sigma^2,c}\, r\, m$ and the deviation from rounding step is bounded, which has probability at least

$$1 - 3\,n^2\,\exp\left(-\text{const}(\sigma^2, c)\, r\, m\right) - 2\,n^2\,\exp\left(-\text{const}(\sigma^2, c)\, r\, m\right) \ge 1 - 5\,n^2\,\exp\left(-\text{const}(\sigma^2, c)\, r\, m\right).$$

Then after allocating all the rest of the items (not in $\{A_i^3\}_{i \in N}$) in an EF1 way, the allocation for this specific ρ guarantees a maximum envy of $1 - K_{\sigma^2,c}\, r\, m/3$. Therefore the algorithm takes the minimum envy over the set of values for ρ, and returns an allocation of envy at most

$$1 - K_{\sigma^2,c}\, r\, m/3 \le 1 - K_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m/3$$

with probability at least

$$1 - 5\,n^2\,\exp\left(-\text{const}(\sigma^2, c)\, r\, m\right) \ge 1 - \exp\Big(-\underbrace{const(\sigma^2, c) \cdot (1 - 1/e)}_{\text{denote this by } E_{\sigma^2,c}} \cdot \min(p, 1/n) \cdot m + 2\,\ln n + \ln 5\Big).$$

Similar to what we did in Section 4.3, we can find range of parameters in both scenarios, such that,

$$E_{\sigma^2,c} \cdot \min(p, 1/n) \cdot m - 2\,\ln n - \ln 5 \longrightarrow \infty$$

for a dominating probability, and

$$\min(p, 1/n) \cdot m \longrightarrow \infty$$

for the envy to vanish.

[9]Our rounding only allows one $X_{j,\alpha}$ among all $j \in N$ to be rounded to 1, which can be achieved by a roulette-like procedure.

In the first scenario, i.e., if $n \to \infty$, $m \to \infty$, $m \geq 3/E_{\sigma^2,c}\, n \ln n$, and $p \geq \frac{3/E_{\sigma^2,c} \ln n}{m}$, then $\min(p, 1/n) \geq \frac{3/E_{\sigma^2,c} \ln n}{m}$, which shows that both terms go to infinity scaling in $\Omega(\ln n)$.

In the second scenario, i.e., if $m \to \infty$, n is bounded above, and $p \in \omega(1/m)$, then $\min(p, 1/n) \cdot m \in \min(\omega(1/m), \Theta(1))\, m \to \infty$, and term still goes to infinity when multiplied by a positive constant and when subtracting $O(1)$ terms from it.

□

F WORST-CASE HARDNESS OF FINDING PROPORTIONAL AND PARETO-OPTIMAL ALLOCATIONS FROM AN ARBITRARY PROPORTIONAL ALLOCATION

Proposition 14. *Suppose there was an algorithm that, when provided with a (worst-case) utility profile and a proportional allocation for that utility profile, would return a proportional* and Pareto-optimal allocation *in polynomial time. Then* P = NP.

Proof. We will prove this by showing a reduction from the NP-complete *partition* problem to this problem. The *partition* problem is the task where given a set of integers $\mathcal{S} = \{p_1, p_2, \ldots, p_k\}$ such that $\sum_{i=1}^{k} p_i = 2K$, decide whether S can be partitioned into two subsets $\mathcal{S}_1, \mathcal{S}_2$ such that the sum of elements in each subset is exactly K.

For the *partition* problem, we can construct a corresponding allocation instance (Table 1): 2 agents and $k + 2$ items, both agent 1 and agent 2 have values p_i $(i = 1, 2, \ldots, k)$ for the first k items. For the last two items, which we denote as α and β, agent 1 has value K for both of them while agent 2 has value $K + 1/2$ for α and $K - 1/2$ for β. Let M denote the set of items. The allocation $\mathcal{A}_0 : A_1^0 = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}, A_2^0 = \{\alpha, \beta\}$ is proportional since $u_1\left(A_1^0\right) = u_2\left(A_2^0\right) = 1/2 \cdot u_1(M) = 1/2 \cdot u_2(M)$. We provide such allocation to the algorithm and the algorithm returns a proportional and Pareto-optimal allocation $\mathcal{A} : A_1, A_2$. We state that $\alpha \in A_2$ and $\beta \in A_1$ if and only if the *partition* problem is a true instance, i.e., an equal partition exists.

$\Rightarrow$: If $\alpha \in A_2$ and $\beta \in A_1$, since $u_i(A_i) \geq 1/2 \cdot u_i(M) = 2K$ for $i = 1, 2$, we have $u_1(A_1/\{\beta\}) \geq K$ and $u_2(A_2/\{\alpha\}) \geq K - 1/2$. For the items in $A_1/\{\beta\}$ and $A_2/\{\alpha\}$ all have integer values, then it can only be that $u_1(A_1/\{\beta\}) = u_2(A_2/\{\alpha\}) = K$. Note that $A_1/\{\beta\}, A_2/\{\alpha\}$ is a partition of the first k items, this result suggests that there is a partition of $\mathcal{S}$ into two subsets that both have summation of K.

$\Leftarrow$: If $\mathcal{S}$ can be partitioned into two subsets that both have summation of K over all elements, then there is an allocation $\mathcal{A}_1 : A_1^1, A_2^1$. It gives the first k items to two agents according to the partition of $\mathcal{S}$, α to agent 2 and β to agent 1. Then this allocation results in $u_1(A_1^1) = K, u_2(A_2^1) = K + 1/2$. Now let's consider the allocation of α and β in $\mathcal{A}$. If α and β are allocated to the same agent, then all the rest of the utilities must go to the other agent so that their utility will be at least K, then this allocation results in utility K for both agents, which is Pareto dominated by our previous $\mathcal{A}_1$, contradicting to the fact that $\mathcal{A}$ is Pareto-optimal. If α goes to agent 1 and β goes to agent 2 in $\mathcal{A}$, then by switching α to agent 2 and β back to agent 1, agent 2 will have 1/2 more utility while agent 1's utility remains the same. Hence the allocation after the switch Pareto dominates $\mathcal{A}$, contradicting to the fact that $\mathcal{A}$ is Pareto-optimal. Therefore it is only possible that α goes to agent 2 and β goes to agent 1 in $\mathcal{A}$.

Hence if exists, the algorithm can solve *partition* problem, from hardness of which we get P = NP.

□

G MIXED MANNA AND MIXED BOOSTS

In this manuscript we assumed that all items are *goods*, meaning that their values are nonnegative. One may also consider *chores*, which are items of negative value (that nevertheless must be allocated). The most general setting (for additive valuation functions) is that of a *mixed manna*, in which some

Agent i	$u_i(\alpha_1)$	$u_i(\alpha_2)$	$\cdots$	$u_i(\alpha_k)$	$u_i(\alpha)$	$u_i(\beta)$
1	p_1	p_2	$\cdots$	p_k	K	K
2	p_1	p_2	$\cdots$	p_k	$K + 1/2$	$K - 1/2$

Table 1. Allocation instance for *partition* problem.

items may be goods and other may be chores (and moreover, the same item might be regarded as a good by one agent and as a chore by another agent). In worst case scenarios, fair allocation in settings of mixed manna is more challenging than fair allocation of goods [11]. However, our results for the smoothed model easily extend to the mixed manna setting (in which the base value of every item is in the range $[-1, 1]$). In fact, no part of our analysis made use of an assumption that item values are non-negative. A minor issues is that, when mixed manna are involved, the allocation produced by round robin might not satisfy EF1 (i.e., agent i might envy agent j even when any one item from agent j's bundle is removed), but satisfies a related property (i does not envy j when one good from j's bundle and one chore from i's bundle are both removed). However, this technical distinction has no effect on our analysis.

As our manuscript focused on goods, it was natural to require that boosts are positive, rather than of mixed sign, because adding a negative boost to a good might create a chore. When moving to mixed manna, it becomes natural to also allow the boosts to be of mixed sign, sometimes positive, sometimes negative. We refer to this setting as that of *mixed boosts*. We now explain how the mixed boosts setting can be reduced to that of positive boosts.

When boosts were positive, the probabilistic model for adding noise to base values involved two parameters, the boost probability p and the boost distribution $\mathcal{D}$. The separation into two parameters was done so as to illustrate that even if p is very small (and hence the boosted valuation function agrees with the base valuation function on most items), the results still hold. However, from a mathematical point of view, the parameter p is redundant. It can be incorporated into the boost distribution $\mathcal{D}$, creating a boost distribution $\mathcal{D}^p$ which is 0 with probability $1 - p$ and distributed like $\mathcal{D}$ with probability p. For goods, we assumed that the boost $\mathcal{D}^p$ is non-negative and supported on $[0, c]$. For mixed manna, we allow mixed boosts, and $\mathcal{D}^p$ is supported on $[-c, c]$. This mixed boosts setting is reducible to the positive boost setting as follows. Subtract c from all base values (we do not care if this causes base values to become negative, as we are dealing with the case of mixed manna), and change the boost distribution to $\mathcal{D}^p_c = \mathcal{D}^p + c$, which is supported on the non-negative interval $[0, 2c]$.

The reduction from mixed boosts to non-negative boosts does not preserve p (the probability of nonzero boost). The probability that $\mathcal{D}^p_c$ outputs a strictly positive boost (which we denote here by p') is the probability that the original distribution $\mathcal{D}^p$ has value strictly larger than $-c$, and not the probability that $\mathcal{D}^p$ has non-zero value. When using our reduction to transfer results from the non-negative boosts case to the mixed boosts case, the parameter p that appeared in the non-negative boosts case is replaced by p'. Unlike p, which is a parameter of intrinsic interest (specifying the probability with which the base value is boosted), the parameter p' does not appear to have intrinsic interest (for most natural choices of mixed sign boost distributions, p' will be close to 1, or even 1. For this reason, when considering mixed boosts, it may be preferable to ignore p' completely, and instead express bounds as functions of other parameters associated with $\mathcal{D}^p_c$. Specifically, our analysis (when viewed this way) shows that the relevant parameters are c (determining the range of the support), the expected highest value among n independent draws from $\mathcal{D}^p_c$, and the expected value of a single draw from $\mathcal{D}^p_c$. The gap between these last two expectations is the expected advantage in boost that the agent with highest boost for an item gets when receiving the item, and this is the quantity that drives the probabilistic analysis in this paper.