

On the Existence of Fair Allocations for Goods and Chores under Dissimilar Preferences

Egor Gagushin*

Marios Mertzani[†]

Alexandros Psomas[‡]

Abstract

We study the fundamental problem of fairly allocating a multiset $\mathcal{M}$ of t types of indivisible items among d groups of agents, where all agents within a group have identical additive valuations. Gorantla et al. [GMV23] showed that for every such instance, there exists a finite number μ such that, if each item type appears at least μ times, an envy-free allocation exists. Their proof is non-constructive and only provides explicit upper bounds on μ for the cases of two groups ($d = 2$) or two item types ($t = 2$).

In this work, we resolve one of the main open questions posed by Gorantla et al. [GMV23] by deriving explicit upper bounds on μ that hold for arbitrary numbers of groups and item types. We introduce a significantly simpler, yet powerful technique that not only yields constructive guarantees for indivisible goods but also extends naturally to chores and continuous domains, leading to new results in related fair division settings such as cake cutting.

Contents

1	Introduction	1
1.1	Our Contributions	1
1.2	Related Work	3
2	Preliminaries	4
3	Conditions Under Which Envy-Free Allocations Exist	5
4	Extensions	9
4.1	Envy-free Allocations of Chores	10
4.2	Envy-free Cake Cutting	12
4.3	Proportional Allocations of Goods and Chores	13
4.3.1	Allocating Goods via the Trading Post Mechanism	13
4.3.2	Proportional Allocations of Goods in the Stochastic Setting	15
4.3.3	Allocating Chores via the Inverse Trading Post Mechanism	15
4.3.4	Proportional Allocations of Chores in the Stochastic Setting	16
A	Proofs missing from Section 3	21
B	Proofs missing from Section 4.1	22
C	Proofs missing from Section 4.2	25
D	Proofs missing from Section 4.3	27
E	Allocating Goods to Similar Agents	31

*Purdue University. eggag03@gmail.com

[†]Purdue University. mmertzan@purdue.edu

[‡]Purdue University. apsomas@purdue.edu

1 Introduction

Fairly allocating resources is a central problem in computer science and economics. What “fairness” means, however, admits multiple formal definitions. A classic notion is *proportionality*: an allocation among n agents is proportional if each agent receives at least a $1/n$ fraction of its total value for all items. Arguably the gold standard is *envy-freeness*, which asks that no agent prefers another agent’s bundle to their own.

Envy-free allocations are not guaranteed to exist in the worst case; consider, for example, a single indivisible item and two agents with a positive value for it. Motivated by such inherent limitations, a substantial body of work tries to identify conditions and settings under which fair allocations become possible. This line of work has culminated in numerous relaxations of envy-freeness, including *envy-freeness up to one item* (EF1) and *envy-freeness up to any item* (EFX); the former is known to exist quite generally, while the latter remains elusive. In stochastic settings envy-free allocations can emerge with high probability; e.g., when agents have additive preferences, drawn i.i.d. from a distribution with a bounded probability density function, Manurangsi and Suksompong [MS21] show that, for the case of goods, an envy-free allocation exists with high probability when the number of items satisfies $m \in \Omega(n \log n / \log \log n)$.

A complementary, structural route asks whether *multiplicity alone* can guarantee fairness. In this direction, Gorantla et al. [GMV23] investigate whether envy-free allocations exist when items are available in multiple identical copies. Formally, they consider the problem of fairly allocating a multiset $\mathcal{M}$ consisting of t types of indivisible items among d groups of agents, where all agents within a group have identical additive valuations. [GMV23] show that, for every such instance, there exists a finite number μ such that, if each item type has at least μ copies, then an envy-free allocation exists. Their proof relies on elegant but intricate arguments demonstrating that a certain linear program has a sufficiently large feasibility space to guarantee the existence of an integral feasible solution.

However, this approach has some key limitations: (i) it does not yield an efficient algorithm for finding such allocations when they are guaranteed to exist, (ii) it is too specialized to easily yield results for other fair-division settings, and, more importantly, (iii) it does not provide explicit bounds on μ , as the argument is purely existential by nature. Gorantla et al. [GMV23] derive upper bounds on μ only for the special cases where either the number of item types is two ($t = 2$) or the number of agent groups is two ($d = 2$).

In this paper, we resolve one of the main questions left open by [GMV23] by providing explicit bounds on μ for arbitrary instances and developing a technique that addresses the aforementioned limitations.

1.1 Our Contributions

We provide the first *explicit* upper bounds on μ , the number of copies required per item type to guarantee the existence of an envy-free allocation. Specifically, we prove that $\mu \in O\left(\frac{tn^3}{g\delta^2}\right)$, where g denotes the greatest common divisor of the agent group sizes, and δ quantifies the degree of dissimilarity between the valuations of different groups. For comparison, Gorantla et al. [GMV23] established that, when $d = 2$, one has $\mu \in O\left(\frac{n^2\sqrt{t}}{g\delta}\right)$ and $\mu \in \Omega\left(\frac{n^2}{g\delta}\right)$, while for the case of $t = 2$ they obtained $\mu \in O\left(\frac{n^2}{g\delta}\right)$. Interestingly, in the special case where each group consists of a single agent, our bound simplifies to $\mu \in O\left(\frac{n^3}{\delta^2}\right)$, showing that, unlike in [GMV23], our result is independent of the number of item types t .

At the heart of our analysis is a dissimilarity condition that implies the existence of envy-free allocations. Our condition formalizes a natural intuition: *fairness should be possible if agents’ preferences are sufficiently dissimilar*. We quantify this dissimilarity between agents’ valuations using well-established measures of distance, including the Euclidean distance and various f -divergences such as the Kullback–Leibler (KL) divergence and the χ^2 -divergence.

Our technique is simple and naturally extends across models, yielding results in a broad range of fair-division settings.

Chores and envy-freeness. We extend our analysis to the setting where items represent *chores*: indivisible tasks that impose costs rather than provide value to agents. By following the same framework developed for the case of goods, we derive explicit upper bounds on μ that guarantee the existence of envy-free allocations in this setting as well.

Cake Cutting. We consider the classic envy-free cake cutting problem, in the query model of Robertson and Webb. It is known that there exists a protocol for finding an envy-free allocation that uses at most $O(n^{n^{n^{n^2}}})$ queries [AM16] (the best known lower bound is $\Omega(n^2)$ [Pro09]). If agents' valuation functions can be represented as polynomials of degree at most d , Brânzei [Brâ15] provides a protocol for finding an envy-free allocation using $O(n^2 d)$ queries. As an application of our condition for the existence of envy-free allocations, we prove that if (i) valuation functions are k -Lipschitz continuous for some constant k , and (ii) the ℓ_2 distance between the valuations of any pair of agents is at least a constant, then we can find a strongly envy-free allocation using $O(n\sqrt{n})$ queries.

Proportionality for Goods. Using the same technical framework, we also derive conditions under which proportional allocations exist for both goods and chores. For proportionality, the intuitive interpretation of our condition is that *each agent must be sufficiently far from society's aggregate valuation*. In the case of goods, society's valuation for an item is defined as the average of the agents' ℓ_1 -normalized valuations, and an agent's distance from society is measured via the χ^2 -divergence between her ℓ_1 -normalized valuation vector and society's valuation vector (see Theorem 7 for the formal statement). In the case of chores, society's cost for an item is given by the harmonic mean of the agents' ℓ_1 -normalized costs.

Leveraging the variational characterization of the χ^2 -divergence, we show that in a stochastic setting, where each value (or cost) is independently drawn from $U[0, 1]$, with high probability all agents are sufficiently far from society when $m \in \Omega(n)$. By combining these results, we recover the state-of-the-art existence guarantees (up to constant factors) for proportional allocations in the stochastic setting.

Our Techniques

The proofs of all aforementioned results follow the same high-level structure. Here, we outline the main arguments for the case of envy-freeness and goods.

First, we identify fractional mechanisms that admit closed-form guarantees on the envy gap: the worst-case margin by which a group values its own bundle over another's. We introduce a new mechanism, the RELATIVE NORM mechanism (and its counterpart, the LOG RELATIVE NORM mechanism, for chores), and show that under this mechanism, the envy gap between agents belonging to different groups is lower bounded by a simple, interpretable quantity that depends on a suitable measure of dissimilarity between the groups.

Second, we formulate a linear program that computes a *fractional* allocation maximizing the minimum envy gap. By definition, the envy gap achieved by the optimal fractional solution is at least as large as that guaranteed by our RELATIVE NORM mechanism. We then develop a rounding procedure that transforms this fractional solution into an integral one while ensuring that (i) each agent receives an integral bundle of items, (ii) all agents within the same group receive identical bundles, and (iii) only a small number of fractional items need to be transferred between groups.

A central technical challenge arises from the requirement that all agents in the same group must receive identical bundles [GMV23]. This constraint complicates the redistribution of fractional items when the total number of copies to be reallocated cannot be expressed as an integer linear combination of group sizes. For example, suppose we need to reassign 11 copies of an item among two groups of sizes $n_1 = 5$ and $n_2 = 7$. In this case, it is impossible to distribute the items while maintaining identical allocations within each group. To overcome this difficulty, we leverage a result of Brauer [Bra42] for the Frobenius coin problem.

Combining the above observations, we show that if the total value of items transferred from one group of agents to another is smaller than the original envy gap, then the resulting integral allocation remains envy-free. The total transferred value can be upper bounded by the product of (i) the total fraction of items exchanged among groups and (ii) the maximum value assigned to any single item. Consequently, we derive a sufficient condition ensuring envy-freeness of the rounded allocation: if the maximum item value is smaller than a threshold that depends on the pairwise distances between agents' valuation vectors, then the final integral allocation is guaranteed to be envy-free.

Finally, we show that as the number of item copies increases, the normalized value of each individual item decreases, while the pairwise distances between agents' valuation vectors remain essentially unchanged. Combining these ingredients, we obtain an explicit bound on the number of copies required to ensure that our condition on the maximum item value is satisfied, guaranteeing the existence of an envy-free allocation.

Future Directions. Beyond the immediate consequences of our work, we believe that our framework provides a promising new lens for understanding the tractability of fairness notions for indivisible items, such as the maximin share guarantee (MMS) and envy-freeness up to any item (EFX). One possibly useful interpretation of our results is that they identify structural features of instances, namely dissimilarity among the agents, that determine when these problems are easy.¹ Using this perspective, our results complement well-known tractable cases for MMS and EFX, such as when agents have identical valuations.

Motivated by this observation, we also explore the problem of allocating indivisible goods to agents that are similar, but not identical. We measure similarity using an f -divergence between normalized valuations. As a proof of concept, in Theorem 11 we prove that when the TV distance between the ℓ_1 -normalized valuations of any pair of agents is upper bounded, then a relaxation of EFX called transfer-EFX (tEFX) can be guaranteed for any number of agents.

1.2 Related Work

The literature on fair allocations of indivisible goods is vast. Beyond the work of [GMV23] that we already discussed, here we briefly mention some other works that are closely related to our paper. We refer the reader to Amanatidis et al. [AAB⁺23] for a recent survey on the allocation of indivisible goods (that also briefly covers some of the results about chores). See Guo et al. [GLD23] for a survey about chores. We also refer the reader to Barnabel [Bar05] for measure-theoretic conditions under which strongly proportional and strongly envy-free allocations exist for goods and chores in a general (cake-cutting) setting.

Relaxations of Proportionality and Envy-Freeness. Faced with the non-existence of proportional and envy-free allocations for indivisible items, a natural approach is to consider relaxed notions of fairness.

A well-studied relaxation of proportionality is maximin share (MMS) fairness [Bud11]. An allocation is MMS if each agent prefers her allocation to the allocation she would have received if she were to split the items into n bundles, and then receives the worst one. Unfortunately, MMS allocations might not exist in the worst case, for goods [KPW18] or chores [ARSW17]. Two well-studied relaxations of envy-freeness are envy-freeness up to one item (EF1) [LMMS04] and envy-freeness up to any item (EFX) [CKM⁺19]. EF1 allocations always exist, even for agents with monotone valuation functions [LMMS04], while EFX allocations are only known to exist for $n \leq 3$ additive agents [CGM24, AAC⁺25]; the case of $n \geq 4$ agents is a major open problem in fair division.

For general additive valuations, a long series of works in fair division strives to pin down the exact constant for which α -MMS allocations exist (for goods, we know that $\alpha \leq \frac{39}{40}$ [FST21] and $\alpha \geq \frac{10}{13}$ [HKSS25]; for chores, we know that $\alpha \leq 13/11$ [HSH23] and $\alpha \geq 44/43$ [FST21]), as well as attack the EFX question from different angles, e.g., approximations (for goods, 0.618-EFX allocations are known to exist [AMN20], and this approximation can be improved in certain cases [AFRS24]; for chores, 4-EFX is known [GMQ25]), or partial allocations (for goods, EFX allocations exist, with at most $n - 2$ unallocated items [Mah24]).

Closer to our work, another major line of work tries to attack the MMS and EFX problems by restricting agents' valuations. For the case of goods, it is known that EFX allocations exist for three types of agents [HGNV25] (this implies existence for identical agents), graphs [CFKS23], lexicographic preferences [HSVX21], bi-valued and dichotomous instances [BEF21, ABFR⁺21, GM23, HPPS20], and two types of items [GMV23]. For the case of chores, EFX allocations are known to exist under leveled preferences [GHLTC23], two types of chores [ALRS23], and binary valuations [TWYZ25]. Hummel [Hum23] shows that, for any integer $c > 0$, there exists a number n_c , such that an MMS allocation exists for any instance with $n \geq n_c$ agents and at most $n + c$ items, where $n_c \leq \lfloor 0.6597^c \cdot c! \rfloor$ for allocations of goods and $n_c \leq \lfloor 0.7838^c \cdot c! \rfloor$ for allocations of chores. Barman et al. [BNV23] prove that MMS allocations for chores exist for binary supermodular cost functions.

Fair allocations in stochastic settings. Dickerson et al. [DGK⁺14] initiated the study of fair division under stochastic preferences, and proved that maximizing utilitarian welfare produces an envy-free allocation with high probability when the number of goods $m \in \Omega(n \log n)$ and items values are drawn i.i.d. from a fixed

¹Every proportional allocation is MMS, and every envy-free allocation is EFX, so our results imply the existence of MMS/EFX as well.

distribution with bounded PDF. Interestingly, [DGK⁺14] implicitly use the existence of a strongly envy-free fractional allocation (the expected allocation of an agent) to get their result; in an online setting, Benadè et al. [BKP⁺24] also use a similar approach.

Manurangsi and Suksompong [MS21] improve the result of [DGK⁺14] and show that an envy-free allocation exists with high probability when $m \in \Omega(n \log n / \log \log n)$. The same work shows that a proportional allocation of goods exists with high probability for $m \geq n$. Manurangsi and Suksompong [MS25] study chore allocation in the i.i.d. setting, and prove that, for PDF bounded distributions, envy-free allocations exist with high probability if $m \geq 2n$, and proportional allocations exist with high probability if $m \in \omega(1)$. Bai and Gözl [BG22] study goods and independent but non-identically distributed additive agents. Benadè et al. [BHPV24] study envy-free allocations of goods for agents with non-additive stochastic valuations. Beyond envy-freeness and proportionality in stochastic settings, Kurokawa et al. [KPW16], Amanatidis et al. [AMNS17], and Farhadi et al. [FGH⁺19] study the existence of MMS allocations (for goods). We note that all aforementioned works treat the distributions as constants (or are for fixed distributions; e.g., the results of Amanatidis et al. [AMNS17] are for the $U[0, 1]$ distribution). Halpern et al. [HPVX25] show that envy-free allocations with high probability exist even without this assumption, and even in an online setting.

Finally, for the case of goods, Bai et al. [BFGP22] study a smoothed model, where worst-case instances are randomly perturbed, and show that envy-free allocations exist with high probability.

2 Preliminaries

There is a set $\mathcal{N}$ of n agents. Each agent i belongs to one of d possible groups: group i has n_i agents that have identical, additive preferences. Without loss of generality, we assume that $n_1 \leq n_2 \leq \dots \leq n_d$. We have that $\sum_{i=1}^d n_i = n$.

Our arguments hinge on the ability to express large integers as nonnegative combinations of the group sizes. Towards this, we rely on the following theorem, due to Brauer [Bra42], and the ensuing corollary.

Theorem 1 ([Bra42]). *Let $n_1 \leq n_2 \leq \dots \leq n_d$, be positive integers such that $\gcd(n_1, \dots, n_d) = 1$. Then, any integer $k \geq (n_1 - 1)(n_d - 1)$ can be written as $k = \sum_{i=1}^d x_i \cdot n_i$, where $x_i \geq 0$ is an integer for all $i \in [d]$.*

Corollary 1. *Let $n_1 \leq n_2 \leq \dots \leq n_d$, be positive integers such that $\gcd(n_1, \dots, n_d) = g$. Then, any integer $k \geq g \cdot (n_1/g - 1) \cdot (n_d/g - 1)$, such that $k \equiv 0 \pmod{g}$, can be written as $k = \sum_{i=1}^d x_i \cdot n_i$, where $x_i \geq 0$ is an integer for all $i \in [d]$.*

For the remainder of this paper, we will use g for the greatest common divisor of $n_1, \dots, n_d$, and θ for the threshold from Corollary 1.

Definition 1 ($\gcd g$ and threshold θ). *Let $g := \gcd(n_1, \dots, n_d)$, and $\theta := g \left(\frac{n_1}{g} - 1 \right) \left(\frac{n_d}{g} - 1 \right)$.*

There is also a multiset $\mathcal{M}$ of m items. There are t types of items, with $k_j \geq 1$ items of type j . We have that $\sum_{j=1}^t k_j = m$. We denote by $\mathbf{k} = [k_1, \dots, k_t]^T$ the vector of the number of copies for each item type. The practical motivation for this problem is the fair allocation of food donations [AAGW15, MPV24, LKK⁺19], where there are many types of items, corresponding to different brands, packages, and so on, and for each type there are many identical copies.

The items can be either *all goods* or *all chores*. For the case of goods, each group of agents $i \in [d]$ has a vector of values $v_i(\mathbf{k}) \in \mathbb{R}^m$ where for each $j \in \mathcal{M}$, $v_{i,j}(\mathbf{k}) \geq 0$ denotes the value that each agent in group i has for item j . The first k_1 entries of $v_i(\mathbf{k})$ are all equal to the value of an agent in group i for the first type of items, the next k_2 entries of $v_i(\mathbf{k})$ are all equal to the value for the second type of items, and so on. The value of an agent in group i for a subset of items $S \subseteq \mathcal{M}$ is $v_i(\mathbf{k}, S) = \sum_{j \in S} v_{i,j}(\mathbf{k})$. For the case where $\mathbf{k} = \mathbf{1}$ (i.e., all types of items have one copy), we drop the parentheses and let $v_i(\mathbf{1}) = v_i$.

For the case of chores, each group of agents $i \in [d]$ has a vector of costs $c_i(\mathbf{k}) \in \mathbb{R}^m$ where for each $j \in \mathcal{M}$, $c_{i,j}(\mathbf{k}) > 0$ denotes the cost that each agent in group i has for receiving item j . The first k_1 entries of $c_i(\mathbf{k})$ are all equal to the cost for the first type of items, the next k_2 entries of $c_i(\mathbf{k})$ are all equal to the cost for the second type of items, and so on. The cost of an agent in group i for a subset of items $S \subseteq \mathcal{M}$ is $c_i(\mathbf{k}, S) = \sum_{j \in S} c_{i,j}(\mathbf{k})$. Similarly we let $c_i(\mathbf{1}) = c_i$.

Throughout the paper, it will be convenient to work with normalized preferences. For ease of notation, we use $\tilde{v}_i^{(1)}(\mathbf{k}) \in [0, 1]^m$ to refer to the ℓ_1 -normalized valuation of an agent in group i , that is, $\tilde{v}_{i,j}^{(1)}(\mathbf{k}) = \frac{v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k}, \mathcal{M})\|_1} = \frac{v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k})\|_1}$ for $j \in \mathcal{M}$. We use notation $\tilde{v}_i^{(2)}(\mathbf{k})$ to refer to the ℓ_2 -normalized valuation of an agent in group i , that is, $\tilde{v}_{i,j}^{(2)}(\mathbf{k}) = \frac{v_{i,j}(\mathbf{k})}{\sqrt{\sum_{j \in \mathcal{M}} (v_{i,j}(\mathbf{k}))^2}} = \frac{v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k})\|_2}$ for $j \in \mathcal{M}$. Similarly, $\tilde{c}_i^{(1)}(\mathbf{k}) \in [0, 1]^m$ is defined as $\tilde{c}_{i,j}^{(1)}(\mathbf{k}) = \frac{c_{i,j}(\mathbf{k})}{c_i(\mathbf{k}, \mathcal{M})} = \frac{c_{i,j}(\mathbf{k})}{\|c_i(\mathbf{k})\|_1}$, for all $j \in \mathcal{M}$; we will not use ℓ_2 -normalized costs (and therefore skip the corresponding notation).

We write $v_{\min}(\mathbf{k})$ (resp. $c_{\min}(\mathbf{k})$) and $v_{\max}(\mathbf{k})$ (resp. $c_{\max}(\mathbf{k})$) for the minimum and maximum value (resp. cost) of an item. That is, $v_{\min}(\mathbf{k}) = \min_{i \in [d], j \in \mathcal{M}} v_{i,j}(\mathbf{k})$, $v_{\max}(\mathbf{k}) = \max_{i \in [d], j \in \mathcal{M}} v_{i,j}(\mathbf{k})$, $c_{\min}(\mathbf{k}) = \min_{i \in [d], j \in \mathcal{M}} c_{i,j}(\mathbf{k})$, and $c_{\max}(\mathbf{k}) = \max_{i \in [d], j \in \mathcal{M}} c_{i,j}(\mathbf{k})$. The ℓ_1 -normalized and ℓ_2 -normalized minimum and maximum values/costs, $\tilde{v}_{\max}^{(1)}(\mathbf{k})$, $\tilde{v}_{\min}^{(1)}(\mathbf{k})$, $\tilde{c}_{\max}^{(1)}(\mathbf{k})$ and so on, are defined analogously.

Allocations. A fractional allocation $x \in [0, 1]^{d \cdot m}$ defines, for every agent group $i \in [d]$ and item $j \in \mathcal{M}$, the fraction $x_{i,j} \geq 0$ of the item j that each agent in group i receives. An allocation x is feasible if, for all items $j \in \mathcal{M}$, $\sum_{i \in [d]} n_i \cdot x_{i,j} \leq 1$ and $x_{i,j} \geq 0$ for all $i \in [d]$, $j \in \mathcal{M}$. Throughout the paper, unless explicitly stated, all agents in the same group will receive the exact same bundle of items, and therefore (for the purposes of discussing envy-freeness) it suffices to refer to groups instead of individual agents. We are interested in *integral* allocations, i.e., each item can be allocated to at most one agent. For ease of notation, we use the notation A for integral allocations and x for fractional allocations. An integral allocation $A = (A_1, \dots, A_d)$ is a partition of $\mathcal{M}$, where A_i denotes the set of items allocated to each agent in group $i \in [d]$, and $A_{i,z} \in \mathbb{Z}_{\geq 0}$ indicates the number of copies each agent in group i received from item type z . Throughout the paper, we only consider *complete* allocations: an allocation is complete if every item gets allocated.

Fairness Notions An (integral or fractional) allocation of goods x is *proportional* if, for all groups $i \in [d]$, $v_i(\mathbf{k}, x_i) \geq n_i \cdot \frac{v_i(\mathbf{k}, \mathcal{M})}{n}$; if the inequality is strict, then the allocation is called *strongly proportional*. An (integral or fractional) allocation of goods x is *envy-free* if, for all pairs of groups $i, i' \in \mathcal{N}$, $v_i(\mathbf{k}, x_i) \geq v_i(\mathbf{k}, x_{i'})$; if the inequality is strict, then the allocation is called *strongly envy-free*. Let $\text{GAP}_{i,i'}(x) = v_i(x_i) - v_i(x_{i'})$ be the advantage (or “negative envy”) of an agent in group i for an agent in group i' , in an allocation x of goods.

An (integral or fractional) allocation of chores x is *proportional* if, for all agents $i \in \mathcal{N}$, $c_i(x_i) \leq \frac{c_i(\mathcal{M})}{n}$; if the inequality is strict, then the allocation is called *strongly proportional*. An (integral or fractional) allocation of chores x is *envy-free* if, for all pairs of agents $i, i' \in \mathcal{N}$, $c_i(x_i) \geq c_i(x_{i'})$; if the inequality is strict, then the allocation is called *strongly envy-free*. For chores, we similarly define $\text{CGAP}_{i,i'}(x) = c_i(x_{i'}) - c_i(x_i)$ as the advantage of an agent in group i for an agent in group i' in allocation x .

Statistical Distances An f -divergence $D_f(P||Q)$ measures the distance between two probability measures P and Q . For the most part, in this paper, we will consider f -divergences for probability mass functions (i.e., discrete distributions), and specifically, we will be studying the f -divergence between ℓ_1 -normalized preference vectors.

Let $P = (p_1, \dots, p_z)$ and $Q = (q_1, \dots, q_z)$ be two probability mass functions ($p_i, q_i \geq 0$ for all i , and $\sum_{i=1}^z p_i = \sum_{i=1}^z q_i = 1$). The χ^2 -divergence between P and Q is defined as $D_{\chi^2}(P||Q) := \sum_{i=1}^z \frac{(P_i - Q_i)^2}{Q_i}$. The KL-divergence between P and Q is defined as $D_{KL}(P||Q) := \sum_{i=1}^z P_i \ln \left(\frac{P_i}{Q_i} \right)$. Finally, the Total variation (TV) distance between P and Q is defined as $D_{TV}(P||Q) := \frac{1}{2} \sum_{i=1}^z |P_i - Q_i|$.

3 Conditions Under Which Envy-Free Allocations Exist

In this section, we study the existence of envy-free allocations. Missing proofs are deferred to Section A.

We start by defining a mechanism for constructing a good *fractional* allocation: the RELATIVE NORM mechanism, given valuations $v_1, \dots, v_d$, outputs an allocation x^{RN} such that

$$x_{i,j}^{\text{RN}} = \frac{\tilde{v}_{i,j}^{(2)}(\mathbf{k})}{n \tilde{v}_{\max}^{(2)}(\mathbf{k})} + \frac{1}{n} - \frac{\sum_{i' \in [d]} n_{i'} \cdot \tilde{v}_{i',j}^{(2)}(\mathbf{k})}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})}$$

where $x_{i,j}^{\text{RN}}$ is the fraction of item $j \in \mathcal{M}$ that each agent of group i receives, $i \in [d]$. Recall that $\tilde{v}_{i,j}^{(2)}(\mathbf{k})$ is the ℓ_2 -normalized value of an agent in group i for item j , and $\tilde{v}_{\max}^{(2)}(\mathbf{k})$ is the ℓ_2 -normalized maximum value.

Although the above mechanism may initially appear complicated or arbitrary, it is in fact derived from simple and natural principles. Our goal is to design a mechanism in which the fraction of an item allocated to an agent increases with how much the agent values that item. Since our target property (i.e. envy-freeness) is scale-free, we must employ a normalized notion of value so that agents with uniformly higher valuations across all items do not gain an unfair advantage. So, the most natural idea is to let the allocation scale linearly with the agent's normalized valuation, i.e., $x_{i,j} = a \cdot \tilde{v}_{i,j}^{(2)}(\mathbf{k}) + b$ for some choice of a, b . The parameters a and b are chosen to ensure that all fractional allocations are nonnegative and that every item is fully allocated. Following this reasoning, we obtain our RELATIVE NORM mechanism.

First, we argue that this mechanism is feasible.

Lemma 1. *The RELATIVE NORM mechanism is feasible. That is, (i) $x_{i,j}^{\text{RN}} \geq 0$ for all $i \in [d]$ and $j \in \mathcal{M}$, and (ii) $\sum_{i \in \mathcal{N}} n_i \cdot x_{i,j}^{\text{RN}} = 1$ for all $j \in \mathcal{M}$.*

Proof. Towards proving that $x_{i,j}^{\text{RN}} \geq 0$ we have:

$$\begin{aligned} x_{i,j}^{\text{RN}} &= \frac{\tilde{v}_{i,j}^{(2)}(\mathbf{k})}{n \tilde{v}_{\max}^{(2)}(\mathbf{k})} + \frac{1}{n} - \frac{\sum_{i' \in [d]} n_{i'} \cdot \tilde{v}_{i',j}^{(2)}(\mathbf{k})}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \\ &\geq \frac{1}{n} - \frac{1}{n^2} \sum_{i' \in [d]} n_{i'} \cdot \frac{\tilde{v}_{i',j}^{(2)}(\mathbf{k})}{\tilde{v}_{\max}^{(2)}(\mathbf{k})} \\ &\geq \frac{1}{n} - \frac{1}{n^2} \sum_{i' \in [d]} n_{i'} \\ &= \frac{1}{n} - \frac{1}{n} = 0 \end{aligned}$$

Towards proving that $\sum_{i \in [d]} n_i \cdot x_{i,j}^{\text{RN}} = 1$ we have:

$$\begin{aligned} \sum_{i \in [d]} n_i \cdot x_{i,j}^{\text{RN}} &= \sum_{i \in [d]} n_i \cdot \left(\frac{\tilde{v}_{i,j}^{(2)}(\mathbf{k})}{n \tilde{v}_{\max}^{(2)}(\mathbf{k})} + \frac{1}{n} - \frac{\sum_{i' \in [d]} n_{i'} \cdot \tilde{v}_{i',j}^{(2)}(\mathbf{k})}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \right) \\ &= \sum_{i \in [d]} n_i \cdot \left(\frac{1}{n} + \frac{\sum_{i' \in [d]} n_{i'} (\tilde{v}_{i,j}^{(2)}(\mathbf{k}) - \tilde{v}_{i',j}^{(2)}(\mathbf{k}))}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \right) \\ &= 1 + \underbrace{\frac{1}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \sum_{i \in [d]} \sum_{i' \in [d]} n_i \cdot n_{i'} \cdot (\tilde{v}_{i,j}^{(2)}(\mathbf{k}) - \tilde{v}_{i',j}^{(2)}(\mathbf{k}))}_0 \\ &= 1. \end{aligned}$$

□

Let $\text{GAP}_{i,i'}(x) := v_i(x_i) - v_i(x_{i'})$ denote the advantage (or, equivalently, negative envy) of an agent of group i for an agent of group i' under allocation x . The next lemma gives a lower bound on $\text{GAP}_{i,i'}(x^{\text{RN}})$.

Lemma 2. *For all $i, i' \in [d]$, $\text{GAP}_{i,i'}(x^{\text{RN}}) \geq \frac{\|v_i(\mathbf{k})\|_2}{2n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \cdot \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2$.*

Proof.

$$\text{GAP}_{i,i'}(x^{\text{RN}}) = \sum_{j \in \mathcal{M}} v_{i,j}(\mathbf{k})(x_{i,j}^{\text{RN}} - x_{i',j}^{\text{RN}})$$

$$\begin{aligned}
&= \|v_i(\mathbf{k})\|_2 \sum_{j \in \mathcal{M}} \frac{v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k})\|_2} \left(\left(\frac{\tilde{v}_{i,j}^{(2)}(\mathbf{k})}{n \tilde{v}_{\max}^{(2)}(\mathbf{k})} + \frac{1}{n} - \frac{\sum_{i'' \in [d]} n_{i''} \cdot \tilde{v}_{i'',j}^{(2)}(\mathbf{k})}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \right) \right. \\
&\quad \left. - \left(\frac{\tilde{v}_{i',j}^{(2)}(\mathbf{k})}{n \tilde{v}_{\max}^{(2)}(\mathbf{k})} + \frac{1}{n} - \frac{\sum_{i'' \in [d]} n_{i''} \cdot \tilde{v}_{i'',j}^{(2)}(\mathbf{k})}{n^2 \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \right) \right) \\
&= \|v_i(\mathbf{k})\|_2 \sum_{j \in \mathcal{M}} \tilde{v}_{i,j}^{(2)}(\mathbf{k}) \left(\frac{\tilde{v}_{i,j}^{(2)}(\mathbf{k}) - \tilde{v}_{i',j}^{(2)}(\mathbf{k})}{n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \right) \\
&= \frac{\|v_i(\mathbf{k})\|_2}{n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \sum_{j \in \mathcal{M}} \left(\left(\tilde{v}_{i,j}^{(2)}(\mathbf{k}) \right)^2 - \tilde{v}_{i,j}^{(2)}(\mathbf{k}) \cdot \tilde{v}_{i',j}^{(2)}(\mathbf{k}) \right) \\
&= \frac{\|v_i(\mathbf{k})\|_2}{n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} \sum_{j \in \mathcal{M}} \left(\frac{1}{2} (\tilde{v}_{i,j}^{(2)}(\mathbf{k}))^2 + \frac{1}{2} (\tilde{v}_{i',j}^{(2)}(\mathbf{k}))^2 - \tilde{v}_{i,j}^{(2)}(\mathbf{k}) \cdot \tilde{v}_{i',j}^{(2)}(\mathbf{k}) \right) \\
&= \frac{\|v_i(\mathbf{k})\|_2 \cdot \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2}{2n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})},
\end{aligned}$$

where in the penultimate equation we used the fact that $\sum_{j \in \mathcal{M}} k_j (\tilde{v}_{i,j}^{(2)}(\mathbf{k}))^2 = \sum_{j \in \mathcal{M}} k_j (\tilde{v}_{i',j}^{(2)}(\mathbf{k}))^2 = 1$. $\square$

Next, consider the following linear program, which finds the fractional allocation that maximizes the minimum $\text{GAP}_{i,i'}$, for ℓ_2 -normalized valuations. If an allocation x has a non-negative objective value with respect to this LP, then x is envy-free.

$$\begin{aligned}
&\text{maximize} && \alpha \\
&\text{subject to} && \alpha \leq \sum_{j \in \mathcal{M}} \tilde{v}_{i,j}^{(2)}(\mathbf{k}) \cdot (x_{i,j} - x_{i',j}) \quad \forall i, i' \in [d], i' \neq i \\
& && \sum_{i=1}^d n_i \cdot x_{i,j} \leq 1 \quad \forall j \in \mathcal{M} \\
& && \alpha, x_{i,j} \geq 0 \quad \forall i \in [d], \forall j \in \mathcal{M}
\end{aligned} \tag{1}$$

The next lemma argues about the existence of an integral allocation whose minimum GAP is not too far from the optimal value of LP (1).

Lemma 3. *If $k_z \geq \theta$ and $k_z \equiv 0 \pmod{g}$ for all $z \in [t]$ then there exists an integral allocation A such that, for all $i \in [d]$, $\min_{i' \in [d]} \text{GAP}_{i,i'}(A) \geq \|v_i(\mathbf{k})\|_2 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} v_{i,j}(\mathbf{k})$, where α^* is the optimal value of LP (1).*

Proof. LP (1) has $dm + 1$ variables (α and the $x_{i,j}$ s), therefore, there exists an optimal solution with $dm + 1$ tight constraints. There are $d(d-1) + m + dm + 1$ constraints, out of which dm constraints are non-negativity constraints for $x_{i,j}$. By simple counting, we have that at least $dm + 1 - (d(d-1) + m + 1)$ out of these non-negativity constraints must be tight. Therefore, at most $m + d(d-1)$ of the “ $x_{i,j} \geq 0$ ” are *not* tight. Since, without loss of generality, $x_{i,j} > 0$ for some $i \in [d]$ and all $j \in \mathcal{M}$ (otherwise, we could set $x_{i,j} = 1/n$ for all $i \in [d]$ without affecting the objective), at most $d(d-1)$ items are shared between groups of agents.

Let x^* be an optimal solution of LP (1) where at most $d(d-1)$ items are shared between groups. For each group i and item type $z \in [t]$, let $B_{i,z} = \sum_{j \in \mathcal{M}: j \text{ has type } z} x_{i,j}^*$ be the total fraction of type- z items that group i received under x^* . Our aim is to make $B_{i,z}$ s integral, without changing them too much.

For each $z \in [t]$, initialize $S_z \leftarrow 0$. We will have three phases of decreasing $B_{i,z}$ s and increasing the corresponding S_z . After these phases, we will guarantee that for every $i \in [d]$ and $z \in [t]$, $B_{i,z} \equiv 0 \pmod{n_i}$. Therefore, a $B_{i,z}$ fraction of items of type z can be integrally and equally allocated to the individual agents of each group i , for every $i \in [d]$. We will also guarantee that $S_z = \sum_{i \in [d]} y_{i,z} \cdot n_i$ for some set of non-negative integers $\{y_{i,z}\}_{i \in [d]}$; this implies that a S_z fraction of items of type z can be integrally distributed among groups in a way that, for each group i , individuals of group i receive the same number of items.

Phase 1. Consider every item $\ell \in \mathcal{M}$ such that there exist $i, i' \in [d]$ (with $i \neq i'$) with $x_{i,\ell}^* > 0$ and $x_{i',\ell}^* > 0$; that is, ℓ is an item that is fractionally shared among at least two groups of agents. For every such item ℓ , set $B_{i,\ell} \leftarrow B_{i,\ell} - x_{i,\ell}^*$, for all $i \in [d]$, and $S_z \leftarrow S_z + \sum_{i \in [d]} x_{i,\ell}^*$, where z is the type of item ℓ . From our discussion above, we have that there are at most $d(d-1)$ such items ℓ . Therefore, at the end of Phase 1, we have that $\sum_{z \in [t]} S_z \leq d(d-1)$.

Phase 2. At the start of Phase 2, we have that for every item $j \in \mathcal{M}$ of type z , $B_{i,z}$ is an integer, for all $i \in [d]$. Let $B_{i,z} = n_i \cdot q_{i,z} + r_{i,z}$, where $r_{i,z} \in [0, n_i - 1]$. For every $z \in [t]$, set $B_{i,z} \leftarrow B_{i,z} - r_{i,z}$ for every $i \in [d]$, and set $S_z \leftarrow S_z + \sum_{i \in [d]} r_{i,z}$. Therefore, at the end of Phase 2, we have that $B_{i,z} = n_i \cdot q_{i,z}$ for some integer $q_{i,z}$, and that $\sum_{z \in [t]} S_z \leq d(d-1) + t \sum_{i \in [d]} (n_i - 1) = d(d-1) + t(n-d)$.

Phase 3. For every $z \in [t]$ while $S_z < \theta$, we pick an arbitrary group of agents $i \in [d]$, such that $B_{i,z} > 0$, and set $B_{i,z} \leftarrow B_{i,z} - n_i$ and $S_z \leftarrow S_z + n_i$. First, note that it cannot be that $B_{i,z} = 0$ for all $i \in [d]$ and $S_z < \theta$, since $k_z = \sum_{i \in [d]} B_{i,z} + S_z \geq \theta$; therefore, if $S_z < \theta$ then $B_{i,z} > 0$. Moreover, at the end of Phase 2, we are guaranteed that $B_{i,z} = n_i \cdot q_{i,z}$, for some non-negative integer $q_{i,z}$; therefore $B_{i,z} \leftarrow B_{i,z} - n_i$ is a valid operation. Repeating until $S_z \geq \theta$ for all $z \in [t]$, we have that $\sum_{z \in [t]} S_z \leq d(d-1) + t(n-d) + t(\theta + n_d - 1)$. We also have that for each $z \in [t]$, $\sum_{i \in [d]} B_{i,z} = \sum_{i \in [d]} n_i \cdot q'_{i,z}$ for some set of integers $\{q'_{i,z}\}_{i \in [d]}$. Thus, $\sum_{i \in [d]} B_{i,z} \equiv 0 \pmod{g}$. But since $k_z \equiv 0 \pmod{g}$ and $S_z + \sum_{i \in [d]} B_{i,z} = k_z$, we also have that $S_z \equiv 0 \pmod{g}$ for all $z \in [t]$.

Since $S_z \equiv 0 \pmod{g}$ and $S_z \geq \theta$, applying *Corollary 1* we have that there exists a set of non-negative integers $\{y_{i,z}\}_{i \in [d]}$ such that $S_z = \sum_{i \in [d]} y_{i,z} \cdot n_i$. We can therefore allocate $y_{i,z} \cdot n_i$ items of type z to each group i , for all $i \in [d]$, $z \in [t]$. Since $\sum_{i \in [d]} B_{i,z} = \sum_{i \in [d]} n_i \cdot q'_{i,z}$ for some set of integers $\{q'_{i,z}\}_{i \in [d]}$, we can also allocate $q'_{i,z}$ items of type z to each group i . Let A be this allocation, and notice that A allocates all items. We have that $\sum_{z \in [t]} y_{i,z} \leq \sum_{z \in [t]} S_z \leq d(d-1) + t(n-d) + t(\theta + n_d - 1)$. Thus, after Phase 3, every agent group has received at most $d(d-1) + t(n-d) + t(\theta + n_d - 1)$ extra items (compared to the allocation x^*). All agents of the same group have received exactly the same integral bundle. Also, compared to the allocation x^* , each agent in group i has lost at most $(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} v_{i,j}(\mathbf{k})$ value and each agent in any other group has gained at most that amount in value. Thus $\min_{i' \in [d]} \text{GAP}_{i,i'}(A) \geq \min_{i' \in [d]} \text{GAP}_{i,i'}(x^*) - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} v_{i,j}(\mathbf{k}) = \|v_i(\mathbf{k})\|_2 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} v_{i,j}(\mathbf{k})$. $\square$

Given these ingredients, we can prove that an envy-free allocation exists if we have enough items of each type, and the maximum ℓ_2 -normalized values are small.

Theorem 2. *If $k_z \geq \theta$ and $k_z \equiv 0 \pmod{g}$, for all $z \in [t]$, and $\tilde{v}_{\max}^{(2)}(\mathbf{k}) \leq \sqrt{\frac{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2}{2n(d(d-1) + t(\theta + n + n_d - d - 1))}}$, then an envy-free allocation exists.*

Proof. Let α^* be the optimal value of LP (1); Lemma 3 implies that there exists an integral solution A such that, for all $i \in [d]$, $\min_{i' \in [d]} \text{GAP}_{i,i'}(A) \geq \|v_i(\mathbf{k})\|_2 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in \mathcal{M}} v_{i,j}(\mathbf{k})$. x^{RN} is a feasible solution to LP (1), therefore, $\alpha^* \geq \min_{i, \tilde{i} \in [d]} \frac{\text{GAP}_{i,\tilde{i}}(x^{\text{RN}})}{\|v_i(\mathbf{k})\|_2}$, which, by Lemma 2, is at least $\min_{i, \tilde{i} \in [d]} \frac{\|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{\tilde{i}}^{(2)}(\mathbf{k})\|_2^2}{2n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})}$. Putting everything together, we have that for all $i \in [d]$, $\min_{i' \in [d]} \text{GAP}_{i,i'}(A) \geq \|v_i(\mathbf{k})\|_2 \min_{i, \tilde{i} \in \mathcal{N}} \frac{\|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{\tilde{i}}^{(2)}(\mathbf{k})\|_2^2}{2n \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})} - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in \mathcal{M}} v_{i,j}$. Therefore, if, for all $i \in [d]$, $\max_{j \in \mathcal{M}} \tilde{v}_{i,j}^{(2)}(\mathbf{k}) \leq \min_{i, \tilde{i} \in \mathcal{N}} \frac{\|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{\tilde{i}}^{(2)}(\mathbf{k})\|_2^2}{2n(d(d-1) + t(\theta + n + n_d - d - 1)) \cdot \tilde{v}_{\max}^{(2)}(\mathbf{k})}$, then A is envy-free (since all $\text{GAP}_{i,i'}(A)$ are positive). Re-arranging and simplifying, we get the theorem. $\square$

As a corollary to Theorem 2, we can get the following simpler expression for the case where every group has only one agent (i.e., $d = n$, or $n_i = 1$ for all i), by observing that $\theta = 0$ in this case.

Corollary 2. *If $d = n$ and $\tilde{v}_{\max}^{(2)}(\mathbf{k}) \leq \sqrt{\frac{\min_{i, i' \in [d]} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2}{2n^3}}$, then an envy-free allocation exists.*

In the remainder of this section, we aim to prove a version of Theorem 2 that removes the second condition on the maximum ℓ_2 -normalized value, and only lower bounds the number of items of each type.

We start with a couple of technical lemmas.

Lemma 4. *For any pair $i, i' \in [d]$, $\|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2 \geq \sqrt{\frac{\alpha}{\beta}} \cdot \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 - \left(1 - \sqrt{\frac{\alpha}{\beta}}\right)$ where $k_z \in [\alpha, \beta]$ for each $z \in [t]$.*

Lemma 5. *$x \left(\sqrt{\frac{x}{x+a}} \cdot b - \left(1 - \sqrt{\frac{x}{x+a}}\right) \right)^2 \geq b^2 x - a(b+1)$, where $a, b \geq 0$ and $x \geq \frac{a(b+1)}{2b}$.*

We are now ready to prove our main result.

Theorem 3. *If $k_z \geq \frac{2n(d^2+t(\theta+n+n_d-d-1))}{\min_{i,i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2}$ and $k_z \equiv 0 \pmod{g}$ for all $z \in [t]$, an envy-free allocation exists.*

Proof. Let $\mu = \frac{2n(d^2+t(\theta+n+n_d-d-1))}{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2}$. First, notice that, without loss of generality, we can assume that $k_z \in [\mu, \mu + n - 1]$. If that is not the case, and $k_z > \mu + n - 1$, we can simply allocate one copy of item z to each agent and continue this process until the condition is satisfied; since all agents receive the exact same amount of items, envy is not affected by this process. Also, since we allocate copies in multiples of n and $n \equiv 0 \pmod{g}$, then the remaining number of copies is also divisible by g (so, the condition $k_z \equiv 0 \pmod{g}$ also remains true).

From Theorem 2 we know that if $\tilde{v}_{\max}^{(2)}(\mathbf{k}) \leq \sqrt{\frac{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2}{2n(d(d-1)+t(\theta+n+n_d-d-1))}}$, then an envy-free allocation exists. Notice that:

$$\tilde{v}_{\max}^{(2)}(\mathbf{k}) = \max_{i \in [d]} \frac{\max_{j \in \mathcal{M}} v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k})\|_2} = \max_{i \in [d]} \frac{\max_{z \in [t]} v_{i,z}}{\|v_i(\mathbf{k})\|_2} \leq \max_{i \in [d]} \frac{\max_{z \in [t]} v_{i,z}}{\sqrt{\mu} \|v_i\|_2} \leq \frac{\tilde{v}_{\max}^{(2)}}{\sqrt{\mu}} \leq \frac{1}{\sqrt{\mu}}.$$

Thus, if $\mu \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2 \geq 2n(d(d-1) + t(\theta + n + n_d - d - 1))$ an envy-free allocation exists. We also have that for every $i, i' \in [d]$:

$$\begin{aligned} 0 &= \frac{2n(d^2 + t(\theta + n + n_d - d - 1))}{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2} \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 - 2n(d^2 + t(\theta + n + n_d - d - 1)) \\ &= \mu \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 - 2n(d^2 + t(\theta + n + n_d - d - 1)) \\ &\leq \mu \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 - 2n(d(d-1) + t(\theta + n + n_d - d - 1)) - 4n \quad (d \geq 2) \\ &\leq \mu \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 - 2n(d(d-1) + t(\theta + n + n_d - d - 1)) - (n-1) \left(\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 + 1 \right) \\ &\quad (\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 \leq 2, \text{ using the triangle inequality}) \\ &\leq \mu \left(\sqrt{\frac{\mu}{\mu + n - 1}} \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 - \left(1 - \sqrt{\frac{\mu}{\mu + n - 1}} \right) \right)^2 - 2n(d(d-1) + t(\theta + n + n_d - d - 1)). \\ &\quad (\text{Lemma 5, } \mu \geq \frac{(n-1) \cdot (\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 + 1)}{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2}) \\ &\leq \mu \cdot \min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2^2 - 2n(d(d-1) + t(\theta + n + n_d - d - 1)). \quad (\text{Lemma 4}) \end{aligned}$$

Combining the above, we get that an envy-free allocation exists. $\square$

As an immediate corollary, we can get the following, simpler bound for the case of $d = n$ (i.e., if there is exactly one agent in each group).

Corollary 3. *If $d = n$, and $k_z \geq \frac{2n^3}{\min_{i \in [d], i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2}$ for all $z \in [t]$ then an envy-free allocation exists.*

Theorem 3 implies that there exists a $\mu \in O(tn^3/(g\eta))$, where $\min_{i, i' \in [d]} \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 \geq \eta$, such that if $k_z \geq \mu$, and $k_z \equiv 0 \pmod{g}$, for all $z \in [t]$, then an envy-free allocation exists. [GMV23] prove that, for $d = 2$, $\mu \in O(n^2\sqrt{t}/(g\delta))$ and $\mu \in O(n^2/(g\delta))$, when the two valuations functions are δ -far from being identical; this means that the angle ρ between the two valuations is at least δ . For $t = 2$, [GMV23] prove that $\mu \in O(n^2/(g\delta))$. Using the fact that $\|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2^2 = \|\tilde{v}_i^{(2)}\|_2^2 + \|\tilde{v}_{i'}^{(2)}\|_2^2 - 2\|\tilde{v}_i^{(2)}\|_2 \cdot \|\tilde{v}_{i'}^{(2)}\|_2 \cdot \cos \delta = 2 - 2\cos \delta \geq \frac{4}{\pi^2} \delta^2$ we can get that our results imply that $\mu \in O(tn^3/(g\delta^2))$ for arbitrary d, t . Interestingly, in Corollary 3 we see that $\mu \in O(n^3/\eta)$ (or $\mu \in O(n^3/\delta^2)$) using the δ -far notion of [GMV23], so there is no dependence on the number of item types t .

4 Extensions

In this section, we study various extensions of our technique to related problems in fair division.

4.1 Envy-free Allocations of Chores

In this section, we study the problem of allocating a set of m chores to d groups of agents with additive preferences. Missing proofs can be found in Section B.

We define the LOG-RELATIVE NORM mechanism, which, given costs $c_1(\mathbf{k}), \dots, c_d(\mathbf{k})$ outputs an allocation x^{LRN} such that every agent of group $i \in [d]$ and item $j \in \mathcal{M}$ we have

$$x_{i,j}^{\text{LRN}} = \frac{\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} + \frac{1}{n} - \frac{\prod_{i' \in [d]} \ln \left(\tilde{c}_{i',j}^{(1)}(\mathbf{k}) \right)^{n_{i'}}}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})}.$$

Notice that this is the same mechanism as the RELATIVE NORM mechanism as defined in Section 3, but instead of $\tilde{v}^{(2)}$ we use $-\ln \tilde{c}^{(1)}$ as a proxy for an agent's disutility for a chore. We start by showing that this mechanism outputs a feasible allocation.

Lemma 6. *The LOG-RELATIVE NORM mechanism is feasible: $x_{i,j}^{\text{LRN}} \geq 0$ for all $i \in [d]$ and $j \in \mathcal{M}$, and $\sum_{i \in [d]} n_i \cdot x_{i,j}^{\text{LRN}} = 1$ for all $j \in \mathcal{M}$.*

Proof.

$$\begin{aligned} x_{i,j}^{\text{LRN}} &= \frac{\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} + \frac{1}{n} - \frac{\prod_{i' \in [d]} \ln \left(\tilde{c}_{i',j}^{(1)}(\mathbf{k}) \right)^{n_{i'}}}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \\ &\geq \frac{1}{n} - \frac{1}{n^2} \sum_{i' \in [d]} n_{i'} \frac{\ln \tilde{c}_{i',j}^{(1)}(\mathbf{k})}{\ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \quad \left(\frac{\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \geq 0 \right) \\ &\geq \frac{1}{n} - \frac{\sum_{i' \in [d]} n_{i'}}{n^2} = 0 \quad (|\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})| \leq |\ln \tilde{c}_{\min}^{(1)}(\mathbf{k})|) \\ \sum_{i \in [d]} n_i \cdot x_{i,j}^{\text{LRN}} &= \sum_{i \in [d]} n_i \cdot \left(\frac{\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} + \frac{1}{n} - \frac{\prod_{i' \in [d]} \ln \left(\tilde{c}_{i',j}^{(1)}(\mathbf{k}) \right)^{n_{i'}}}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \right) \\ &= \frac{\sum_{i \in [d]} n_i}{n} + \sum_{i \in [d]} n_i \left(\frac{\sum_{i' \in [d]} n_{i'} \ln \left(\tilde{c}_{i,j}^{(1)}(\mathbf{k}) \right)}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} - \frac{\sum_{i' \in [d]} n_{i'} \ln \left(\tilde{c}_{i',j}^{(1)}(\mathbf{k}) \right)}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \right) \\ &= 1 + \underbrace{\frac{1}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \sum_{i \in [d]} \sum_{i' \in [d]} n_i n_{i'} \left(\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k}) - \ln \tilde{c}_{i',j}^{(1)}(\mathbf{k}) \right)}_0 = 1 \end{aligned}$$

□

Recall that $\text{CGAP}_{i,i'}(x) = c_i(x'_i) - c_i(x_i)$ is the advantage of an agent in i for an agent in i' in allocation x . The next lemma bounds $\text{CGAP}_{i,i'}(x^{\text{LRN}})$ as a function of the KL-divergence between $\tilde{c}_i^{(1)}$ and $\tilde{c}_{i'}^{(1)}$.

Lemma 7. *For all $i, i' \in [d]$, $\text{CGAP}_{i,i'}(x^{\text{LRN}}) = \frac{\|c_i(\mathbf{k})\|_1}{-n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_{i'}^{(1)}(\mathbf{k})).$*

Proof.

$$\begin{aligned} \text{CGAP}_{i,i'}(x^{\text{LRN}}) &= \sum_{j \in \mathcal{M}} c_{i,j}(\mathbf{k}) (x_{i',j}^{\text{LRN}} - x_{i,j}^{\text{LRN}}) \\ &= \|c_i(\mathbf{k})\|_1 \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \left(\left(\frac{\ln \tilde{c}_{i,j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} + \frac{1}{n} - \frac{\prod_{i'' \in [d]} \ln \left(\tilde{c}_{i'',j}^{(1)}(\mathbf{k}) \right)^{n_{i''}}}{n^2 \ln \tilde{c}_{\min}^{(1)}(\mathbf{k})} \right) \right. \end{aligned}$$

$$\begin{aligned}
& - \left(\frac{\ln \tilde{c}_{i',j}^{(1)}(\mathbf{k})}{n \ln \tilde{c}_{min}^{(1)}(\mathbf{k})} + \frac{1}{n} - \frac{\prod_{i'' \in [d]} \ln \left(\tilde{c}_{i'',j}^{(1)}(\mathbf{k}) \right)^{n_{i''}}}{n^2 \ln \tilde{c}_{min}^{(1)}(\mathbf{k})} \right) \\
& = \frac{\|c_i(\mathbf{k})\|_1}{n \ln \tilde{c}_{min}^{(1)}(\mathbf{k})} \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{\tilde{c}_{i,j}^{(1)}(\mathbf{k})}{\tilde{c}_{i',j}^{(1)}(\mathbf{k})} \right) \\
& = \frac{\|c_i(\mathbf{k})\|_1}{-n \ln \tilde{c}_{min}^{(1)}(\mathbf{k})} D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_{i'}^{(1)}(\mathbf{k})).
\end{aligned}$$

□

Similarly to previous results in this paper, we next consider a linear program for maximizing CGAP, this time for ℓ_1 -normalized costs, and then argue that this LP has a good integral solution.

$$\begin{aligned}
& \text{maximize} && \alpha \\
& \text{subject to} && \alpha \leq \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) (x_{i',j} - x_{i,j}) && \forall i, i' \in [d], i' \neq i \\
& && \sum_{i=1}^n n_i \cdot x_{i,j} \leq 1 && \forall j \in \mathcal{M} \\
& && \alpha, x_{i,j} \geq 0 && \forall i \in [d], \forall j \in \mathcal{M}
\end{aligned} \tag{2}$$

Lemma 8. *If $k_z \geq \theta$ and $k_z \equiv 0 \pmod{g}$ for all $z \in [t]$ then there exists an integral allocation A such that, for all $i \in [d]$, $\min_{i' \in [d]} \text{CGAP}_{i,i'}(A) \geq \|c_i(\mathbf{k})\|_1 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} c_{i,j}(\mathbf{k})$, where α^* is the optimal value of LP (2).*

The proof of the lemma is almost identical to that of Lemma 3 and is deferred to the appendix.

Theorem 4. *If $k_z \geq \theta$, $k_z \equiv 0 \pmod{g}$ for all $z \in [t]$, and $\tilde{c}_{max}^{(1)}(\mathbf{k}) \leq \frac{\min_{\tilde{i}, \tilde{i} \in [d]} D_{KL}(\tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}) \parallel \tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}))}{2n(d(d-1) + t(\theta + n + n_d - d - 1)) \ln \tilde{c}_{min}^{(1)}(\mathbf{k})^{-1}}$, then an envy-free allocation exists.*

Proof. Let α^* be the optimal value of LP (2); Lemma 8 implies that there exists an integral solution A such that, for all $i \in [d]$, $\min_{i' \in [d]} \text{CGAP}_{i,i'}(A) > \|c_i(\mathbf{k})\|_1 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in \mathcal{M}} c_{i,j}(\mathbf{k})$. x^{LRN} is a feasible solution to LP (2), therefore, $\alpha^* \geq \min_{\tilde{i}, \tilde{i} \in [d]} \frac{\text{CGAP}_{\tilde{i}, \tilde{i}}(x^{\text{LRN}})}{\|c_{\tilde{i}}(\mathbf{k})\|_1}$, which, by Lemma 7, is at least $\min_{\tilde{i}, \tilde{i} \in [d]} \frac{D_{KL}(\tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}) \parallel \tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}))}{-n \ln \tilde{c}_{min}^{(1)}(\mathbf{k})}$. Putting everything together, we have that for all $i \in [d]$, $\min_{i' \in [d]} \text{CGAP}_{i,i'}(A) \geq \|c_i(\mathbf{k})\|_1 \min_{\tilde{i}, \tilde{i} \in [d]} \frac{D_{KL}(\tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}) \parallel \tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}))}{-n \ln \tilde{c}_{min}^{(1)}(\mathbf{k})} - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in \mathcal{M}} c_{i,j}$. Therefore, if, for all $i \in [d]$, $\max_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \leq \min_{\tilde{i}, \tilde{i} \in [d]} \frac{D_{KL}(\tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}) \parallel \tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}))}{2n(d(d-1) + t(\theta + n + n_d - d - 1)) \ln \tilde{c}_{min}^{(1)}(\mathbf{k})^{-1}}$, then A is envy-free (since all $\text{GAP}_{i,i'}(A)$ are non-negative). □

Corollary 4. *If $d = n$ and $\tilde{c}_{max}^{(1)}(\mathbf{k}) \leq \min_{\tilde{i}, \tilde{i} \in [d]} \frac{D_{KL}(\tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}) \parallel \tilde{c}_{\tilde{i}}^{(1)}(\mathbf{k}))}{2n^3 \ln \tilde{c}_{min}^{(1)}(\mathbf{k})^{-1}}$, then an envy-free allocation exists.*

Before proving the main result of this section, we prove two useful technical lemmas.

Lemma 9. *For any pair $i, i' \in [d]$, $D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_{i'}^{(1)}(\mathbf{k})) \geq \frac{\alpha}{\beta} D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_{i'}^{(1)}(\mathbf{k})) - \left(\ln \left(\frac{\beta}{\alpha} \right) + \frac{\beta - \alpha}{\alpha} \ln \left(\tilde{c}_{min}^{(1)} \right)^{-1} \right)$ where $k_z \in [\alpha, \beta]$ for each $z \in [t]$.*

Lemma 10.

$$x \left(\frac{x}{x+a} b - \ln \frac{x+a}{x} \right) - c \ln(x+a) \geq \frac{b}{2} x - \left(a \left(\frac{3}{2} b + 1 \right) + c \left(\ln \left(\frac{2c}{b} \right) - 1 \right) \right)$$

where $x, a, b, c \geq 0$

We are now ready to prove the main theorem for this section: a condition on the number of copies needed from each item to guarantee that an envy-free allocation for chores exists. The proof follows the same structure as the proof of Theorem 3 and is deferred to the appendix.

Theorem 5. If $k_z \geq 2 \frac{n + (\frac{5}{2}n + \lambda - 1) \ln(\tilde{c}_{min}^{(1)})^{-1} + \lambda \left(\ln \left(\frac{2\lambda}{\min_{i, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)} \parallel \tilde{c}_{\tilde{i}}^{(1)})} \right) - 1 \right)}{\min_{i, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)} \parallel \tilde{c}_{\tilde{i}}^{(1)})}$ and $k_z \equiv 0 \pmod{g}$ for all $z \in [t]$, an envy-free allocation exists, where $\lambda = 2n(d(d-1) + t(\theta + n + n_d - d - 1))$.

Theorem 5 implies that there exists a $\mu \in O \left(\frac{tn^3 \left(\ln(\tilde{c}_{min}^{(1)})^{-1} + \ln \left(\frac{tn^3}{g\eta} \right) \right)}{g\eta} \right)$, where $\min_{i, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)} \parallel \tilde{c}_{\tilde{i}}^{(1)}) \geq \eta$, such that if $k_z \geq \mu$, and $k_z \equiv 0 \pmod{g}$, for all $z \in [t]$, then an envy-free allocation exists.

4.2 Envy-free Cake Cutting

In this section, we consider the classic problem of cake-cutting in the query model of Robertson and Webb.

In this setting, each group has one agent ($n_1 = \dots = n_d = 1$) and every agent $i \in \mathcal{N}$ has a valuation function U_i that assigns a non-negative value to any subinterval of $[0, 1]$. We can think of these values as induced by a density u_i , i.e., for an interval $I = [a, b]$, $U_i(I) = \int_{x \in I} u_i(x) dx$. In the standard model, valuations are normalized, i.e., $U_i([0, 1]) = \int_{x \in [0, 1]} u_i(x) dx = 1$, additive, i.e., for a set of disjoint intervals $I_1, \dots, I_z$, $U_i(\cup_{j=1}^z I_j) = \sum_{j=1}^z U_i(I_j)$, and non-atomic, i.e., $\forall x, y \in [0, 1], \lambda \in [0, 1], \exists z \in [x, y]$ where $U_i([x, z]) = \lambda \cdot U_i([x, y])$. We will additionally assume that (i) for all $i \in \mathcal{N}$, u_i is k -Lipschitz continuous, and (ii) for all $i, i' \in \mathcal{N}$, $\|u_i - u_{i'}\|_2 \geq \sqrt{\delta}$, where $\|u_i\|_2 = \sqrt{\int_{x \in [0, 1]} u_i(x)^2 dx}$ and δ is a positive constant.

In the Robertson and Webb [RW98] model (henceforth, RW) there are two types of queries: (i) $\text{EVAL}_i(I)$: given an agent i and an interval $I \subseteq [0, 1]$, return $U_i(I)$, and (ii) $\text{CUT}_i(x, z)$: given an agent i , a point $x \in [0, 1]$, and a value $z \in [0, 1]$, return the smallest $y \in [x, 1]$ such that $U_i([x, y]) = z$.

It will be convenient to work with $\tilde{u}_i^{(2)}(x) = \frac{u_i(x)}{\|u_i\|_2}$. By definition, $\int_{x \in [0, 1]} \tilde{u}_i^{(2)}(x)^2 dx = 1$. Furthermore, $\tilde{u}_i^{(2)}$ is $\frac{k}{\|u_i\|_2}$ -Lipschitz continuous; since $\|u_i\|_2 \geq 1$ (by the Cauchy-Schwarz inequality), $\tilde{u}_i^{(2)}$ is k -Lipschitz continuous, for all $i \in \mathcal{N}$. Similarly, for all $i, i' \in \mathcal{N}$, $\|\tilde{u}_i^{(2)} - \tilde{u}_{i'}^{(2)}\|_2 \geq \sqrt{\delta}$.

Lemma 11. For all $i \in \mathcal{N}$, $\tilde{u}_i^{(2)}(x) \leq \max\{(8k)^{1/3}, 2\}$, for all $x \in [0, 1]$.

Proof. Let $z^* = \max_{x \in [0, 1]} \tilde{u}_i^{(2)}(x)$, and let x^* be such that $\tilde{u}_i^{(2)}(x^*) = z^*$. Then, $\tilde{u}_i^{(2)}(x) \geq z^* - k|x - x^*|$. Wlog, $x^* \leq 1/2$; let $I = [x^*, x^* + 1/2]$. For all $x \in I$, it holds that $\tilde{u}_i^{(2)}(x) \geq z^* - \frac{k}{2}$. We have that $1 = \int_{x \in [0, 1]} \tilde{u}_i^{(2)}(x)^2 dx \geq \int_{x \in I} (z^* - \frac{k}{2})^2 dx = \frac{1}{2}(z^* - \frac{k}{2})^2$. Re-arranging, we have that $z^* \leq \frac{k}{2} + \sqrt{2}$. For $k \in [1, 2\sqrt{2}]$, it is easy to confirm that $\frac{k}{2} + \sqrt{2} \leq (8k)^{1/3}$, and for $k \leq 1$, we have that $z^* \leq \frac{1}{2} + \sqrt{2} < 2$.

Otherwise, if $k > 2\sqrt{2}$, we have that $\frac{z^*}{2k} \leq \frac{1}{4} + \frac{\sqrt{2}}{2k} \leq \frac{1}{2}$. Then, $x^* + \frac{z^*}{2k} \leq 1$. Let $I' = [x^*, x^* + \frac{z^*}{2k}]$. Repeating the same argument as above, we have that for all $x \in I'$, $\tilde{u}_i^{(2)}(x) \geq z^* - k \frac{z^*}{2k} = \frac{z^*}{2}$. We have $1 = \int_{x \in [0, 1]} \tilde{u}_i^{(2)}(x)^2 dx \geq \int_{x \in I'} (\frac{z^*}{2})^2 dx = \frac{z^*}{2k} \cdot \frac{(z^*)^2}{4} = \frac{(z^*)^3}{8k}$. Re-arranging, we have that $z^* \leq (8k)^{1/3}$. $\square$

We are ready to prove our main result for this section.

Theorem 6. If (i) for all $i \in \mathcal{N}$, u_i is k -Lipschitz continuous, and (ii) for all $i, i' \in \mathcal{N}$, $\|u_i - u_{i'}\|_2 \geq \sqrt{\delta}$, then we can find a strongly envy-free allocation using at most $\max\{nk, \frac{4n^4(\max\{(8k)^{1/3}, 2\})^2}{\sqrt{(3.5k)^2 + 16n^3(\max\{(8k)^{1/3}, 2\})^2\delta} - 3.5k}}\}$ EVAL queries (and no CUT queries), in the RW model.

Proof. First, we can cut the cake into $\frac{1}{\epsilon}$ pieces, each of length ϵ . Lemma 11 implies that the ℓ_2 normalized value of each agent, for any piece, is at most $\epsilon \max\{(8k)^{1/3}, 2\}$. We can find the exact value of each agent for each piece using n/ϵ EVAL queries; we will prove that, under the specified condition on the valuations, there is a strongly envy-free allocation of the pieces.

We have assumed a lower bound on $\int_0^1 (\tilde{u}_i^{(2)}(x) - \tilde{u}_{i'}^{(2)}(x))^2 dx$, the squared ℓ_2 distance between any pair of agents (recall that $\|\tilde{u}_i^{(2)} - \tilde{u}_{i'}^{(2)}\|_2 \geq \sqrt{\delta}$). We would like to bound the difference of $\int_0^1 (\tilde{u}_i^{(2)}(x) - \tilde{u}_{i'}^{(2)}(x))^2 dx$ with its “discretized” version (w.r.t. the pieces we created). Specifically, we have that $\int_{\epsilon j}^{\epsilon(j+1)} \tilde{u}_i^{(2)}(x) dx$ is the real (ℓ_2 -normalized) value of agent i for piece j . Treating pieces as indivisible items, $\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}$

is the ℓ_2 norm of the (ℓ_2 -normalized) valuation vector of agent i ; importantly, this is no longer equal to 1, and therefore we must normalize again! Let $\frac{\int_{\epsilon_j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon_{j'}}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}}$ be the normalized value of agent i

for “item” j . Finally, let $G_{i,i'} = \sum_{j=0}^{\frac{1}{\epsilon}-1} \left(\frac{\int_{\epsilon_j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon_{j'}}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}} - \frac{\int_{\epsilon_j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon_{j'}}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} \right)^2$ be the squared ℓ_2 distance between two valuation vectors with respect to the *indivisible* items/pieces created.

To apply Corollary 2 we need a bound on $\min_{i,i' \in \mathcal{N}} G_{i,i'}$. The following lemma, whose proof is deferred to Section C, gives us our desired bound.

Lemma 12. *For all $1/k \geq \epsilon > 0$, and $i, i' \in \mathcal{N}$,*

$$G_{i,i'} \geq \int_0^1 (\tilde{u}_i^{(2)}(x) - \tilde{u}_{i'}^{(2)}(x))^2 dx - \frac{7k\epsilon}{2}.$$

Lemma 12 implies that $\sqrt{\min_{i,i' \in \mathcal{N}} G_{i,i'}} \geq \sqrt{\delta - 3.5k\epsilon}$, for $\epsilon \leq 1/k$. Therefore, by Corollary 2, we have that if $\epsilon \max\{(8k)^{1/3}, 2\} \leq \frac{\sqrt{\delta - 3.5k\epsilon}}{2n\sqrt{n}}$, or, equivalently, if $\epsilon \leq \min\{\frac{1}{k}, \frac{\sqrt{(3.5k)^2 + 16n^3(\max\{(8k)^{1/3}, 2\})^2 \delta - 3.5k}}{4n^3(\max\{(8k)^{1/3}, 2\})^2}\}$, then a strongly envy-free allocation (of “whole” pieces) exists. $\square$

By treating k and δ as constants, we obtain the following corollary, which provides a more intuitive bound on the query complexity.

Corollary 5 (of Theorem 6). *If (i) for all $i \in \mathcal{N}$, u_i is k -Lipschitz continuous, and (ii) for all $i, i' \in \mathcal{N}$, $\|u_i - u_{i'}\|_2 \geq \sqrt{\delta}$, for some constants k and δ , then we can find a strongly envy-free allocation in the RW model using $O(n\sqrt{n})$ EVAL queries (and no CUT queries).*

4.3 Proportional Allocations of Goods and Chores

In this section, we study conditions under which proportional allocations exist for goods and chores. We also give applications of our existence results.

We will no longer assume that agents are organized into groups. This relaxation does not make the setting any more restrictive, as our bounds will not depend on the pairwise distances between agents. Consequently, the results remain meaningful even when different agents have identical or nearly identical valuations. Moreover, since proportionality is a significantly weaker notion of fairness than envy-freeness, we will not require that agents with identical valuations receive identical bundles. Treating each agent as an individual group therefore simplifies our exposition.

Throughout this section, we also assume that $\mathbf{k} = \mathbf{1}$. While it is possible to replicate the analysis and derive slightly stronger bounds on the number of copies required to guarantee proportionality (as opposed to envy-freeness), such an analysis would closely mirror that of the previous sections and would not yield additional technical insights. Instead, our focus is on exploring how the connection between fair allocations and f -divergences can be leveraged to obtain results by drawing on the heavy machinery developed in information theory for these divergences. In particular, we employ the variational characterization of f -divergences to establish bounds that we then use to recover, up to constant factors, state-of-the-art results in stochastic fair division.

All missing proofs are deferred to Section D.

4.3.1 Allocating Goods via the Trading Post Mechanism

We define $S_j = \frac{1}{n} \sum_{i \in \mathcal{N}} \frac{v_{i,j}}{v_i(\mathcal{M})} = \frac{1}{n} \sum_{i \in \mathcal{N}} \tilde{v}_{i,j}^{(1)}$ as society’s valuation for item $j \in \mathcal{M}$. Equivalently, S_j is the average of agents’ ℓ_1 -normalized valuations. The TRADING POST mechanism outputs an allocation x^{TP} such that $x_{i,j}^{\text{TP}} = \frac{\tilde{v}_{i,j}^{(1)}}{\sum_{i' \in \mathcal{N}} \tilde{v}_{i',j}^{(1)}}$, for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$.

We start by observing that the utility of an agent in the Trading Post mechanism has a nice form, as a function of the χ^2 -divergence between her valuation and society’s valuation.

Lemma 13. For all $i \in \mathcal{N}$, $\frac{v_i(x_i^{\text{TP}})}{v_i(\mathcal{M})} = \frac{1}{n} \left(1 + D_{\chi^2}(\tilde{v}_i^{(1)} \| S) \right)$.

Proof.

$$\begin{aligned}
\frac{v_i(x_i^{\text{TP}})}{v_i(\mathcal{M})} &= \frac{1}{v_i(\mathcal{M})} \sum_{j \in [m]} v_{i,j} \frac{\tilde{v}_{i,j}^{(1)}}{\sum_{i' \in \mathcal{N}} \tilde{v}_{i',j}^{(1)}} \\
&= \sum_{j \in [m]} \frac{(\tilde{v}_{i,j}^{(1)})^2}{\sum_{i' \in \mathcal{N}} \tilde{v}_{i',j}^{(1)}} \\
&= \frac{1}{n} \left(\sum_{j \in [m]} \frac{(\tilde{v}_{i,j}^{(1)})^2}{\frac{1}{n} \sum_{i' \in \mathcal{N}} \tilde{v}_{i',j}^{(1)}} \right) \\
&= \frac{1}{n} \left(1 + \sum_{j \in [m]} \frac{(\tilde{v}_{i,j}^{(1)})^2}{S_j} - 2\tilde{v}_{i,j}^{(1)} + S_j \right) \quad (\|\tilde{v}_i^{(1)}\|_1 = \|S\|_1 = 1) \\
&= \frac{1}{n} \left(1 + \sum_{j \in [m]} \frac{(\tilde{v}_{i,j}^{(1)})^2 - 2\tilde{v}_{i,j}^{(1)} S_j + S_j^2}{S_j} \right) \\
&= \frac{1}{n} \left(1 + \sum_{j \in [m]} \frac{(\tilde{v}_{i,j}^{(1)} - S_j)^2}{S_j} \right) \\
&= \frac{1}{n} \left(1 + D_{\chi^2}(\tilde{v}_i^{(1)} \| S) \right).
\end{aligned}$$

□

Now, consider the following LP that maximizes the ℓ_1 -normalized minimum utility of an agent. An allocation is proportional if its value with respect to this LP is at least $1/n$.

$$\begin{aligned}
&\text{maximize} && \alpha \\
&\text{subject to} && \alpha \leq \sum_{j=1}^m \tilde{v}_{i,j}^{(1)} x_{i,j} \quad \forall i \in \mathcal{N} \\
& && \sum_{i=1}^n x_{i,j} \leq 1 \quad \forall j \in \mathcal{M} \\
& && \alpha, x_{i,j} \geq 0 \quad \forall i \in \mathcal{N}, \forall j \in \mathcal{M}
\end{aligned} \tag{3}$$

The next lemma allows us to argue about the existence of an integral allocation that has high objective value (with respect to the LP above).

Lemma 14. [Essentially Theorem 3 of [EPV24]] Let x^* be an optimal solution to LP (3). There exists an integral allocation A , such that, for all agents $i \in \mathcal{N}$, $v_i(A_i) \geq v_i(x_i^*) - \max_{j \in \mathcal{M}} v_{i,j}$.

Proof. LP (3) maximized the normalized p -mean of utilities, for $p \rightarrow -\infty$. It is known [EPV24] that there exist prices $p_j^* \geq 0$ for each item $j \in \mathcal{M}$, and budgets b_i^* for each agent $i \in \mathcal{N}$, such that (x^*, p^*, b^*) is a Fisher market equilibrium. [BK19] give an algorithm that, given a Fisher market equilibrium (x, p, b) , in polynomial time, find an integral allocation A and budgets b' (that satisfy certain conditions), such that (A, p, b') is a market equilibrium. [EPV24] prove that, for this integral allocation A , for all $i \in \mathcal{N}$, $v_i(A_i) \geq v_i(x_i^*) - v_{i,j}$, for some item j in the maximum bang-per-buck bundle of agent i . The lemma follows. □

Given Lemmas 13 and 14, we are ready to prove our main result in this section.

Theorem 7. Let $D^* = \min_{i \in \mathcal{N}} \left\{ D_{\chi^2}(\tilde{v}_i^{(1)} \| S) \right\}$. If $\tilde{v}_{\max}^{(1)} \leq \frac{D^*}{n}$, then a proportional allocation exists. When the inequality is strict, then a strongly proportional allocation exists.

Proof. Let α^* be the value of an optimal solution to LP (3), and let A be the integral allocation from Lemma 14. We have that $v_i(A_i) \geq \alpha^* v_i(\mathcal{M}) - \max_{j \in \mathcal{M}} v_{i,j}$. Let x^{TP} be the allocation output by the Trading Post mechanism. Since x^{TP} is a feasible solution for LP (3), we have that $\alpha^* \geq \min_{i \in \mathcal{N}} \tilde{v}_i^{(1)}(x_i^{\text{TP}}) = \min_{i \in \mathcal{N}} \frac{v_i(x_i^{\text{TP}})}{v_i(\mathcal{M})}$.

Therefore, using Lemma 13, we have that $\alpha^* \geq \frac{1}{n} (1 + D^*)$, where $D^* = \min_{i \in \mathcal{N}} \{D_{\chi^2}(\tilde{v}_i^{(1)} || s)\}$. Therefore, we overall have $u_i(A) \geq \frac{1}{n} (1 + D^*) v_i(\mathcal{M}) - \max_{j \in \mathcal{M}} v_{i,j} \geq 1/n$, where in the last inequality we used the fact that, for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$, $\frac{v_{i,j}}{v_i(\mathcal{M})} \leq \frac{D^*}{n}$. $\square$

4.3.2 Proportional Allocations of Goods in the Stochastic Setting

As an application of Theorem 7, we consider a setting with stochastic valuations. Specifically, we consider the setting where $v_{i,j}$ is sampled i.i.d. from the $U[0, 1]$ distribution. [MS21] prove that a proportional allocation exists with high probability for $m \geq n$ (their result, in fact, holds as long as the PDF of the distribution is upper and lower bounded by constants). Here, we re-prove this result in a “modular” way, but with a worse dependence on n . Specifically, we show that in this model $\tilde{v}_{max}^{(1)} \leq \frac{\min_{i \in \mathcal{N}} \{D_{\chi^2}(\tilde{v}_i^{(1)} || S)\}}{n}$ holds with high probability, for $m \in \Omega(n)$. This readily implies the existence of a proportional allocation.

Lemma 15. *In the i.i.d. $U[0, 1]$ model ($v_{i,j} \sim U[0, 1]$), for $m \in \Omega(n)$, it holds that $D_{\chi^2}(\tilde{v}_i^{(1)} || S) \geq \frac{n-3}{3n}$, for all $i \in \mathcal{N}$, with high probability $(1 - O(1/n))$.*

Theorem 8. *In the i.i.d. $U[0, 1]$ model ($v_{i,j} \sim U[0, 1]$), a (strongly) proportional allocation exists with high probability $(1 - O(1/n))$, when $m \in \Omega(n)$.*

Proof. Lemma 15 implies that, with high probability, $D^* = \min_{i \in \mathcal{N}} \{D_{\chi^2}(\tilde{v}_i^{(1)} || S)\} \geq \frac{n-3}{3n}$. Also with high probability, for all $i \in \mathcal{N}, j \in \mathcal{M}$, $\frac{v_{i,j}}{v_i(\mathcal{M})} \leq \frac{1}{v_i(\mathcal{M})} \leq \frac{1}{m} \left(\frac{1}{1/2 - \sqrt{\frac{\log n}{m}}} \right)$. Theorem 7 implies that if $\frac{1}{m} \left(\frac{1}{1/2 - \sqrt{\frac{\log n}{m}}} \right) \leq \frac{D^*}{n}$ then a proportional allocation exists. Therefore, if $\frac{1}{m} \left(\frac{1}{1/2 - \sqrt{\frac{\log n}{m}}} \right) \leq \frac{n-3}{3n^2}$ a proportional allocation exists; assuming that $m \geq c'n$, for some positive constant c' , the strict inequality holds for all sufficiently large n . $\square$

4.3.3 Allocating Chores via the Inverse Trading Post Mechanism

Let $H_j = \frac{n}{\sum_{i \in \mathcal{N}} 1/\tilde{c}_{i,j}^{(1)}}$ be society’s cost for item $j \in \mathcal{M}$, defined as the harmonic mean of normalized agent costs. We define a variation of the Trading Post mechanism where weights are inversely proportional to agents’ costs. Formally, given costs $c_1, \dots, c_n$, the INVERSE TRADING POST outputs an allocation $x_{i,j}^{\text{ITP}} = \frac{1/\tilde{c}_{i,j}^{(1)}}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}}$. The cost of agent $i \in \mathcal{N}$ when participating in the inverse Trading Post mechanism is $c_i(x_i^{\text{ITP}}) = \sum_{j \in \mathcal{M}} c_{i,j} \frac{1/\tilde{c}_{i,j}^{(1)}}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}} = \sum_{j \in \mathcal{M}} \frac{c_i(\mathcal{M})}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}} = n c_i(\mathcal{M}) \sum_{j \in \mathcal{M}} H_j$. We start by expressing this cost as a function of the distance between an agent and society.

Lemma 16. *For all $i \in \mathcal{N}$, $\frac{c_i(x_i^{\text{ITP}})}{c_i(\mathcal{M})} = \frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \right)$.*

Interestingly, the expression for the normalized cost of agent i in Lemma 16 does not depend on i ’s identity, i.e., all agents have the same normalized cost.

Next, consider the following LP that minimizes the maximum ℓ_1 -normalized cost of an agent.

$$\begin{aligned} & \text{minimize} && \alpha \\ & \text{subject to} && \alpha \geq \sum_{j=1}^m \tilde{c}_{ij}^{(1)} x_{ij} && \forall i \in \mathcal{N} \\ & && \sum_{i=1}^n x_{ij} = 1 && \forall j \in \mathcal{M} \\ & && \alpha, x_{ij} \geq 0 && \forall i \in \mathcal{N}, \forall j \in \mathcal{M} \end{aligned} \tag{4}$$

Similarly to Lemma 14, the following lemma argues about the existence of a good integral allocation.

Lemma 17. *[Essentially Theorem 11 of [EPV24]] Let x^* be an optimal solution to LP (4). There exists an integral allocation A , such that, for all agents $i \in \mathcal{N}$, $c_i(A_i) \leq c_i(x_i^*) + \max_{j \in \mathcal{M}} c_{i,j}$.*

Proof. Similarly to Lemma 14, we use the market interpretation of LP (4) by [EPV24], coupled with a rounding algorithm of [BS19]. Specifically, LP (4) minimizes the minimizing the normalized p -mean of disutilities, for $p \rightarrow \infty$. Letting x^* be an optimal solution to LP (4), [EPV24] prove that there exist rewards r and budgets e such that (x, r, e) is a market equilibrium. [BS19] give an algorithm (based on the algorithm of [BK19]) that, given such a market equilibrium, computes another market equilibrium (A, r, e') , where A is an integral allocation and e' satisfies certain properties. Relevant to us here, [EPV24] prove that $c_i(A_i) \leq c_i(x_i^*) + \max_{j \in \mathcal{M}} c_{i,j}$; the lemma follows. $\square$

We are now ready to prove the main result of this section.

Theorem 9. *If $\tilde{c}_{max}^{(1)} \leq \frac{1}{n^2} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}}$, then a proportional allocation exists. When the inequality is strict, then a strongly proportional allocation exists.*

Proof. The allocation of Lemma 17 satisfies, for all $i \in \mathcal{N}$, $c_i(A_i) \leq c_i(x_i^*) + \max_{j \in \mathcal{M}} c_{i,j}$. Furthermore, $\frac{c_i(x_i^*)}{c_i(\mathcal{M})} \leq \alpha^*$, where α^* is the optimal value of LP (4). The allocation of the inverse Trading Post mechanism, x^{ITP} , is a feasible solution for LP (4), therefore, $\alpha^* \leq \min_{z \in \mathcal{N}} \frac{c_z(x_z^{\text{ITP}})}{c_z(\mathcal{M})}$. By Lemma 16, the RHS is equal to $\frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \right)$. Combining, we have that

$$\begin{aligned} c_i(A_i) &\leq \frac{c_i(\mathcal{M})}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \right) + \max_{j \in \mathcal{M}} c_{i,j} \\ &\leq \frac{c_i(\mathcal{M})}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \right) + \frac{c_i(\mathcal{M})}{n^2} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \\ &= \frac{c_i(\mathcal{M})}{n}. \end{aligned}$$

$\square$

We note that there exists an equivalent version of this theorem, where the bound on $\tilde{c}_{max}^{(1)}$ involves the χ^2 -divergence between $\tilde{c}_i^{(1)}$ and $H/\|H\|_1$. Specifically, if $\tilde{c}_{max}^{(1)} \leq \frac{1}{n^2} \sum_{\ell \in \mathcal{N}} \left((1 - \|H\|_1)^2 + \|H\|_1^2 D_{\chi^2}(\tilde{c}_\ell^{(1)} \| \frac{H}{\|H\|_1}) \right)$, then a proportional allocation exists.

4.3.4 Proportional Allocations of Chores in the Stochastic Setting

As an application of Theorem 7, similarly to the case of goods, we consider a setting with stochastic preferences: $c_{i,j}$ is sampled i.i.d. from the $U[0, 1]$ distribution. In recent work, [MS25] prove that, for all distributions whose PDF is bounded by constants, a proportional allocation exists with high probability for $m \in \omega(1)$. Here, we re-prove a weaker version of this result, by showing that the condition in Theorem 9 holds with high probability for $m \in \Omega(n)$ (and therefore, by Theorem 9, a proportional allocation exists).

We start by showing a couple of technical lemmas.

Lemma 18. $\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} \geq \langle H - \tilde{c}_i^{(1)}, g \rangle - \frac{\langle \tilde{c}_i^{(1)}, g^2 \rangle}{4}$, where g is any vector in $\mathbb{R}^m$ and $g^2 = [g_1^2, g_2^2, \dots, g_m^2]$.

Lemma 19. *In the i.i.d. $U[0, 1]$ model ($c_{i,j} \sim U[0, 1]$), for $m \in \Omega(n)$, it holds that $\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} \geq \frac{5}{16}$, for all $i \in \mathcal{N}$, with high probability $(1 - O(1/n))$.*

We are now ready to state our result about the existence of proportional allocations of chores.

Theorem 10. *In the i.i.d. $U[0, 1]$ model ($c_{i,j} \sim U[0, 1]$), a (strongly) proportional allocation exists with high probability $(1 - O(1/n))$, when $m \in \Omega(n)$.*

Proof. Lemma 19 gives that, with high probability, $\sum_{j \in \mathcal{M}} \frac{(\bar{c}_{i,j}^{(1)} - H_j)^2}{\bar{c}_{i,j}^{(1)}} \geq \frac{5}{16}$, for all $i \in \mathcal{N}$. Therefore, with high probability, $\sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\bar{c}_{i,j}^{(1)} - H_j)^2}{\bar{c}_{i,j}^{(1)}} \geq \frac{5n}{16}$. Theorem 9 then implies that if, for all $i \in \mathcal{N}$, $\frac{n^2}{c_i(\mathcal{M})} \max_{j \in \mathcal{M}} c_{i,j} \leq \frac{5n}{16}$, or, equivalently, for all $i \in \mathcal{N}, j \in \mathcal{M}$, $\frac{c_{i,j}}{c_i(\mathcal{M})} \leq \frac{5}{16n}$, then a proportional allocation exists. With high probability, we have that for all $i \in \mathcal{N}, j \in \mathcal{M}$, $\frac{c_{i,j}}{c_i(\mathcal{M})} \leq \frac{1}{c_i(\mathcal{M})} \leq \frac{1}{m} \left(\frac{1}{1/2 - \sqrt{\frac{\log n}{m}}} \right)$. Therefore, for $m \geq c'n$, for some constant c' , the desired condition holds for sufficiently large n ; the theorem follows. $\square$

Acknowledgements

Alexandros Psomas and Marios Mertzaniadis are supported in part by an NSF CAREER award CCF-2144208, and a research award from the Herbert Simon Family Foundation.

References

- [AAB⁺23] Georgios Amanatidis, Haris Aziz, Georgios Birmpas, Aris Filos-Ratsikas, Bo Li, Hervé Moulin, Alexandros A Voudouris, and Xiaowei Wu. Fair division of indivisible goods: Recent progress and open questions. *Artificial Intelligence*, 322:103965, 2023.
- [AAC⁺25] Hannaneh Akrami, Noga Alon, Bhaskar Ray Chaudhury, Jugal Garg, Kurt Mehlhorn, and Ruta Mehta. Efx: a simpler approach and an (almost) optimal guarantee via rainbow cycle number. *Operations Research*, 73(2):738–751, 2025.
- [AAGW15] Martin Aleksandrov, Haris Aziz, Serge Gaspers, and Toby Walsh. Online fair division: analysing a food bank problem. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pages 2540–2546, 2015.
- [ABFR⁺21] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, Alexandros Hollender, and Alexandros A Voudouris. Maximum nash welfare and other stories about efx. *Theoretical Computer Science*, 863:69–85, 2021.
- [AFRS24] Georgios Amanatidis, Aris Filos-Ratsikas, and Alkmini Sgouritsa. Pushing the frontier on approximate efx allocations. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 1268–1286, 2024.
- [AGS25] Arash Ashuri, Vasilis Gkatzelis, and Alkmini Sgouritsa. Ef2x exists for four agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 13555–13563, 2025.
- [ALRS23] Haris Aziz, Jeremy Lindsay, Angus Ritossa, and Mashbat Suzuki. Fair allocation of two types of chores. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '23, page 143–151, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems.
- [AM16] Haris Aziz and Simon Mackenzie. A discrete and bounded envy-free cake cutting protocol for any number of agents. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 416–427. IEEE, 2016.
- [AMN20] Georgios Amanatidis, Evangelos Markakis, and Apostolos Ntokos. Multiple birds with one stone: Beating $1/2$ for efx and gmms via envy cycle elimination. *Theoretical Computer Science*, 841:94–109, 2020.
- [AMNS17] Georgios Amanatidis, Evangelos Markakis, Afshin Nikzad, and Amin Saberi. Approximation algorithms for computing maximin share allocations. *ACM Transactions on Algorithms*, 13(4):1–28, 2017.

- [ARSW17] Haris Aziz, Gerhard Rauchecker, Guido Schryen, and Toby Walsh. Algorithms for max-min share fair allocation of indivisible chores. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.
- [Bar05] Julius B Barbanel. *The geometry of efficient fair division*. Cambridge University Press, 2005.
- [BEF21] Moshe Babaioff, Tomer Ezra, and Uriel Feige. Fair and truthful mechanisms for dichotomous valuations. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(6):5119–5126, May 2021.
- [BFGP22] Yushi Bai, Uriel Feige, Paul Gözl, and Ariel D Procaccia. Fair allocations for smoothed utilities. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 436–465, 2022.
- [BG22] Yushi Bai and Paul Gözl. Envy-free and pareto-optimal allocations for agents with asymmetric random valuations. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI*, 2022.
- [BHPV24] Gerdus Benadè, Daniel Halpern, Alexandros Psomas, and Paritosh Verma. On the existence of envy-free allocations beyond additive valuations. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 1287–1287, 2024.
- [BK19] Siddharth Barman and Sanath Kumar Krishnamurthy. On the proximity of markets with integral equilibria. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1748–1755, 2019.
- [BKP⁺24] Gerdus Benade, Aleksandr M Kazachkov, Ariel D Procaccia, Alexandros Psomas, and David Zeng. Fair and efficient online allocations. *Operations Research*, 72(4):1438–1452, 2024.
- [BNV23] Siddharth Barman, Vishnu Narayan, and Paritosh Verma. Fair chore division under binary supermodular costs. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '23*, page 2863–2865, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems.
- [Bra42] Alfred Brauer. On a problem of partitions. *American Journal of Mathematics*, 64(1):299–312, 1942.
- [Brâ15] Simina Brânzei. A note on envy-free cake cutting with polynomial valuations. *Information Processing Letters*, 115(2):93–95, 2015.
- [BS19] Simina Brânzei and Fedor Sandomirskiy. Algorithms for competitive division of chores. *arXiv preprint arXiv:1907.01766*, 2019.
- [Bud11] Eric Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.
- [CFKS23] George Christodoulou, Amos Fiat, Elias Koutsoupias, and Alkmini Sgouritsa. Fair allocation in graphs. In *Proceedings of the 24th ACM Conference on Economics and Computation, EC '23*, page 473–488, New York, NY, USA, 2023. Association for Computing Machinery.
- [CGM24] Bhaskar Ray Chaudhury, Jugal Garg, and Kurt Mehlhorn. Efx exists for three agents. *Journal of the ACM*, 71(1):1–27, 2024.
- [CKM⁺19] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D Procaccia, Nisarg Shah, and Junxing Wang. The unreasonable fairness of maximum nash welfare. *ACM Transactions on Economics and Computation (TEAC)*, 7(3):1–32, 2019.
- [DGK⁺14] John Dickerson, Jonathan Goldman, Jeremy Karp, Ariel Procaccia, and Tuomas Sandholm. The computational rise and fall of fairness. In *Proceedings of the AAAI conference on artificial intelligence*, volume 28, 2014.

- [EPV24] Owen Eckart, Alexandros Psomas, and Paritosh Verma. On the fairness of normalized p-means for allocating goods and chores. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 1267–1267, 2024.
- [FGH⁺19] Alireza Farhadi, Mohammad Ghodsi, Mohammad Taghi Hajiaghayi, Sebastien Lahaie, David Pennock, Masoud Seddighin, Saeed Seddighin, and Hadi Yami. Fair allocation of indivisible goods to asymmetric agents. *Journal of Artificial Intelligence Research*, 64:1–20, 2019.
- [FST21] Uriel Feige, Ariel Sapir, and Laliv Tauber. A tight negative example for mms fair allocations. In *International Conference on Web and Internet Economics*, pages 355–372. Springer, 2021.
- [GHLTC23] Yotam Gafni, Xin Huang, Ron Lavi, and Inbal Talgam-Cohen. Unified fair allocation of goods and chores via copies. *ACM Trans. Econ. Comput.*, 11(3–4), December 2023.
- [GLD23] Hao Guo, Weidong Li, and Bin Deng. A survey on fair allocation of chores. *Mathematics*, 11(16):3616, 2023.
- [GM23] Jugal Garg and Aniket Murhekar. Computing fair and efficient allocations with few utility values. *Theoretical Computer Science*, 962:113932, 2023.
- [GMQ25] Jugal Garg, Aniket Murhekar, and John Qin. Constant-factor efx exists for chores. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1580–1589, 2025.
- [GMV23] Pranay Gorantla, Kunal Marwaha, and Santhoshini Velusamy. Fair allocation of a multiset of indivisible items. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 304–331. SIAM, 2023.
- [HGNV25] Vishwa Prakash Hv, Pratik Ghosal, Prajakta Nimbhorkar, and Nithin Varma. Efx exists for three types of agents. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 101–128, 2025.
- [HKSS25] Ehsan Heidari, Alireza Kaviani, Masoud Seddighin, and AmirMohammad Shahrezaei. Improved maximin share guarantee for additive valuations. *arXiv preprint arXiv:2510.10423*, 2025.
- [HPPS20] Daniel Halpern, Ariel D Procaccia, Alexandros Psomas, and Nisarg Shah. Fair division with binary valuations: One rule to rule them all. In *International Conference on Web and Internet Economics*, pages 370–383. Springer, 2020.
- [HPVX25] Daniel Halpern, Alexandros Psomas, Paritosh Verma, and Daniel Xie. Online envy minimization and multicolor discrepancy: Equivalences and separations. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 188–188, 2025.
- [HSH23] Xin Huang and Erel Segal-Halevi. A reduction from chores allocation to job scheduling. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 908–908, 2023.
- [HSVX21] Hadi Hosseini, Sujoy Sikdar, Rohit Vaish, and Lirong Xia. Fair and efficient allocations under lexicographic preferences. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(6):5472–5480, May 2021.
- [Hum23] Halvard Hummel. On lower bounds for maximin share guarantees. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI ’23*, 2023.
- [KPW16] David Kurokawa, Ariel Procaccia, and Junxing Wang. When can the maximin share guarantee be guaranteed? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.
- [KPW18] David Kurokawa, Ariel D Procaccia, and Junxing Wang. Fair enough: Guaranteeing approximate maximin shares. *Journal of the ACM (JACM)*, 65(2):1–27, 2018.

- [LKK⁺19] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, et al. Webuildai: Participatory framework for algorithmic governance. *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–35, 2019.
- [LMMS04] Richard J Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. On approximately fair allocations of indivisible goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 125–131, 2004.
- [Mah24] Ryoga Mahara. Extension of additive valuations to general valuations on the existence of efx. *Mathematics of operations research*, 49(2):1263–1277, 2024.
- [MPV24] Marios Mertzaniadis, Alexandros Psomas, and Paritosh Verma. Automating food drop: The power of two choices for dynamic and fair food allocation. In *Proceedings of the 25th ACM Conference on Economics and Computation*, EC ’24, page 243, New York, NY, USA, 2024. Association for Computing Machinery.
- [MS21] Pasin Manurangsi and Warut Suksompong. Closing gaps in asymptotic fair division. *SIAM Journal on Discrete Mathematics*, 35(2):668–706, 2021.
- [MS25] Pasin Manurangsi and Warut Suksompong. Asymptotic fair division: Chores are easier than goods. *arXiv preprint arXiv:2504.20704*, 2025.
- [Pro09] Ariel D Procaccia. Thou shalt covet thy neighbor’s cake. In *IJCAI*, pages 239–244, 2009.
- [RW98] Jack Robertson and William Webb. *Cake-cutting algorithms: Be fair if you can*. AK Peters/CRC Press, 1998.
- [TWYZ25] Biaoshuai Tao, Xiaowei Wu, Ziqi Yu, and Shengwei Zhou. On the existence of efx (and pareto-optimal) allocations for binary chores. *Theoretical Computer Science*, 1042:115248, 2025.

A Proofs missing from Section 3

Proof of Lemma 4. First, observe that for every group $\ell \in [d]$:

$$\|v_\ell(\mathbf{k})\|_2 = \sqrt{\sum_{j \in \mathcal{M}} (v_{\ell,j}(\mathbf{k}))^2} = \sqrt{\sum_{z \in [t]} k_z \cdot (v_{\ell,z})^2} \geq \sqrt{\sum_{z \in [t]} \alpha \cdot (v_{\ell,z})^2} = \sqrt{\alpha} \|v_\ell\|_2.$$

Similarly:

$$\|v_\ell(\mathbf{k})\|_2 = \sqrt{\sum_{j \in \mathcal{M}} (v_{\ell,j}(\mathbf{k}))^2} = \sqrt{\sum_{z \in [t]} k_z \cdot (v_{\ell,z})^2} \leq \sqrt{\sum_{z \in [t]} \beta \cdot (v_{\ell,z})^2} = \sqrt{\beta} \|v_\ell\|_2.$$

Thus there exist $\gamma_i, \gamma_{i'}$ such that $\|v_i(\mathbf{k})\|_2 = \frac{\|v_i\|_2}{\gamma_i}$ and $\|v_{i'}(\mathbf{k})\|_2 = \frac{\|v_{i'}\|_2}{\gamma_{i'}}$, such that $\frac{1}{\sqrt{\beta}} \leq \gamma_i \leq \frac{1}{\sqrt{\alpha}}$ and $\frac{1}{\sqrt{\beta}} \leq \gamma_{i'} \leq \frac{1}{\sqrt{\alpha}}$. Without loss of generality, assume that $\gamma_{i'} \geq \gamma_i$. We have that:

$$\begin{aligned} \|\tilde{v}_i^{(2)}(\mathbf{k}) - \tilde{v}_{i'}^{(2)}(\mathbf{k})\|_2 &= \sqrt{\sum_{j \in \mathcal{M}} (\tilde{v}_{i,j}^{(2)}(\mathbf{k}) - \tilde{v}_{i',j}^{(2)}(\mathbf{k}))^2} \\ &= \sqrt{\sum_{j \in \mathcal{M}} \left(\frac{v_{i,j}(\mathbf{k})}{\|v_i(\mathbf{k})\|_2} - \frac{v_{i',j}(\mathbf{k})}{\|v_{i'}(\mathbf{k})\|_2} \right)^2} \\ &= \sqrt{\sum_{j \in \mathcal{M}} \left(\gamma_i \frac{v_{i,j}(\mathbf{k})}{\|v_i\|_2} - \gamma_{i'} \frac{v_{i',j}(\mathbf{k})}{\|v_{i'}\|_2} \right)^2} \\ &= \sqrt{\sum_{j \in \mathcal{M}} \left(\gamma_i \left(\frac{v_{i,j}(\mathbf{k})}{\|v_i\|_2} - \frac{v_{i',j}(\mathbf{k})}{\|v_{i'}\|_2} \right) - (\gamma_{i'} - \gamma_i) \frac{v_{i',j}(\mathbf{k})}{\|v_{i'}\|_2} \right)^2} \\ &\geq \gamma_i \sqrt{\sum_{j \in \mathcal{M}} \left(\frac{v_{i,j}(\mathbf{k})}{\|v_i\|_2} - \frac{v_{i',j}(\mathbf{k})}{\|v_{i'}\|_2} \right)^2} - (\gamma_{i'} - \gamma_i) \sqrt{\sum_{j \in \mathcal{M}} \left(\frac{v_{i',j}(\mathbf{k})}{\|v_{i'}\|_2} \right)^2} \quad (\text{Triangle Inequality}) \\ &= \gamma_i \sqrt{\sum_{z \in [t]} k_z \left(\frac{v_{i,z}}{\|v_i\|_2} - \frac{v_{i',z}}{\|v_{i'}\|_2} \right)^2} - \underbrace{\frac{\gamma_{i'} - \gamma_i}{\gamma_{i'}} \sqrt{\sum_{j \in \mathcal{M}} \left(\frac{v_{i',j}(\mathbf{k})}{\|v_{i'}(\mathbf{k})\|_2} \right)^2}}_1 \\ &\geq \gamma_i \sqrt{\sum_{z \in [t]} \alpha \left(\frac{v_{i,z}}{\|v_i\|_2} - \frac{v_{i',z}}{\|v_{i'}\|_2} \right)^2} - \frac{\gamma_{i'} - \gamma_i}{\gamma_{i'}} \quad (k_z \geq \alpha) \\ &= \gamma_i \sqrt{\alpha} \cdot \sqrt{\sum_{z \in [t]} \left(\frac{v_{i,z}}{\|v_i\|_2} - \frac{v_{i',z}}{\|v_{i'}\|_2} \right)^2} - \frac{\gamma_{i'} - \gamma_i}{\gamma_{i'}} \\ &= \gamma_i \cdot \sqrt{\alpha} \cdot \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 - \frac{\gamma_{i'} - \gamma_i}{\gamma_{i'}} \\ &\geq \sqrt{\frac{\alpha}{\beta}} \cdot \|\tilde{v}_i^{(2)} - \tilde{v}_{i'}^{(2)}\|_2 - \left(1 - \sqrt{\frac{\alpha}{\beta}} \right). \end{aligned}$$

□

Proof of Lemma 5.

$$\begin{aligned} x \left(\sqrt{\frac{x}{x+a}} \cdot b - \left(1 - \sqrt{\frac{x}{x+a}} \right) \right)^2 &= x \left((1 + a/x)^{-1/2} \cdot (b+1) - 1 \right)^2 \\ &\geq x \left(\left(1 - \frac{a}{2x} \right) \cdot (b+1) - 1 \right)^2 \end{aligned}$$

$$\begin{aligned}
&= x \left(b - \frac{a(b+1)}{2x} \right)^2 \\
&\geq x \left(b^2 - \frac{a(b+1)}{x} \right) \\
&= xb^2 - a(b+1)
\end{aligned}$$

where in the second line we used the fact that $(1 + a/x)^{-1/2} \geq 1 - \frac{a}{2x}$ and $x \geq \frac{a(b+1)}{2b}$. $\square$

B Proofs missing from Section 4.1

Proof of Lemma 8. LP (2) has $dm + 1$ variables (α and the $x_{i,j}$ s), therefore, there exists an optimal solution with $dm + 1$ tight constraints. There are $d(d-1) + m + dm + 1$ constraints, out of which dm constraints are non-negativity constraints for $x_{i,j}$. By simple counting, we have that at least $dm + 1 - (d(d-1) + m + 1)$ out of these non-negativity constraints must be tight. Therefore, at most $m + d(d-1)$ of the “ $x_{i,j} \geq 0$ ” are *not* tight. Since, without loss of generality, $x_{i,j} > 0$ for some $i \in [d]$ and all $j \in \mathcal{M}$ (otherwise, we could set $x_{i,j} = 1/n$ for all $i \in [d]$ without affecting the objective), at most $d(d-1)$ chores are shared between groups of agents.

Let x^* be an optimal solution of LP (2) where at most $d(d-1)$ chores are shared between groups. For each group i and item type $z \in [t]$, let $B_{i,z} = \sum_{j \in \mathcal{M}: j \text{ has type } z} x_{i,j}^*$ be the total fraction of type- z item that group i received under x^* . Our aim is to make $B_{i,z}$ s integral, without changing them too much.

For each $z \in [t]$, initialize $S_z \leftarrow 0$. We will have three phases of decreasing $B_{i,z}$ s and increasing the corresponding S_z . After these phases, we will guarantee that for every $i \in [d]$ and $z \in [t]$, $B_{i,z} \equiv 0 \pmod{n_i}$. Therefore, a $B_{i,z}$ fraction of chores of type z can be integrally and equally allocated to the individual agents of each group i , for every $i \in [d]$. We will also guarantee that $S_z = \sum_{i \in [d]} y_{i,z} \cdot n_i$ for some set of non-negative integers $\{y_{i,z}\}_{i \in [d]}$; this implies that a S_z fraction of chores of type z can be integrally distributed among groups in a way that, for each group i , individuals of group i receive the same number of chores.

Phase 1. Consider every item $\ell \in \mathcal{M}$ such that there exist $i, i' \in [d]$ (with $i \neq i'$) with $x_{i,\ell}^* > 0$ and $x_{i',\ell}^* > 0$; that is, ℓ is an item that is fractionally shared among at least two groups of agents. For every such item ℓ , set $B_{i,\ell} \leftarrow B_{i,\ell} - x_{i,\ell}^*$, for all $i \in [d]$, and $S_z \leftarrow S_z + \sum_{i \in [d]} x_{i,\ell}^*$, where z is the type of item ℓ . From our discussion above, we have that there are at most $d(d-1)$ such items ℓ . Therefore, at the end of Phase 1, we have that $\sum_{z \in [t]} S_z \leq d(d-1)$.

Phase 2. At the start of Phase 2, we have that for every item $j \in \mathcal{M}$ of type z , $B_{i,z}$ is an integer, for all $i \in [d]$. Let $B_{i,z} = n_i \cdot q_{i,z} + r_{i,z}$, where $r_{i,z} \in [0, n_i - 1]$. For every $z \in [t]$, set $B_{i,z} \leftarrow B_{i,z} - r_{i,z}$ for every $i \in [d]$, and set $S_z \leftarrow S_z + \sum_{i \in [d]} r_{i,z}$. Therefore, at the end of Phase 2, we have that $B_{i,z} = n_i \cdot q_{i,z}$ for some integer $q_{i,z}$, and that $\sum_{z \in [t]} S_z \leq d(d-1) + t \sum_{i \in [d]} (n_i - 1) = d(d-1) + t(n-d)$.

Phase 3. For every $z \in [t]$ while $S_z < \theta$, we pick an arbitrary group of agents $i \in [d]$, such that $B_{i,z} > 0$, and set $B_{i,z} \leftarrow B_{i,z} - n_i$ and $S_z \leftarrow S_z + n_i$. First, note that it cannot be that $B_{i,z} = 0$ for all $i \in [d]$ and $S_z < \theta$, since $k_z = \sum_{i \in [d]} B_{i,z} + S_z \geq \theta$; therefore, if $S_z < \theta$ then $B_{i,z} > 0$. Moreover, at the end of Phase 2, we are guaranteed that $B_{i,z} = n_i \cdot q_{i,z}$, for some non-negative integer $q_{i,z}$; therefore $B_{i,z} \leftarrow B_{i,z} - n_i$ is a valid operation. Repeating until $S_z \geq \theta$ for all $z \in [t]$, we have that $\sum_{z \in [t]} S_z \leq d(d-1) + t(n-d) + t(\theta + n_d - 1)$. We also have that for each $z \in [t]$, $\sum_{i \in [d]} B_{i,z} = \sum_{i \in [d]} n_i \cdot q'_{i,z}$ for some set of integers $\{q'_{i,z}\}_{i \in [d]}$. Thus, $\sum_{i \in [d]} B_{i,z} \equiv 0 \pmod{g}$. But since $k_z \equiv 0 \pmod{g}$ and $S_z + \sum_{i \in [d]} B_{i,z} = k_z$, we also have that $S_z \equiv 0 \pmod{g}$ for all $z \in [t]$.

Since $S_z \equiv 0 \pmod{g}$ and $S_z \geq \theta$, applying *Corollary 1* we have that there exists a set of non-negative integers $\{y_{i,z}\}_{i \in [d]}$ such that $S_z = \sum_{i \in [d]} y_{i,z} \cdot n_i$. We can therefore allocate $y_{i,z} \cdot n_i$ chores of type z to each group i , for all $i \in [d]$, $z \in [t]$. Since $\sum_{i \in [d]} B_{i,z} = \sum_{i \in [d]} n_i \cdot q'_{i,z}$ for some set of integers $\{q'_{i,z}\}_{i \in [d]}$, we can also allocate $q'_{i,z}$ chores of type z to each group i . Let A be this allocation, and notice that A allocates all chores. We have that $\sum_{z \in [t]} y_{i,z} \leq \sum_{z \in [t]} S_z \leq d(d-1) + t(n-d) + t(\theta + n_d - 1)$. Thus, after Phase 3, every agent group has received at most $d(d-1) + t(n-d) + t(\theta + n_d - 1)$ extra chores (compared to the allocation x^*). All agents of the same group have received exactly the same integral bundle. Also, compared to the allocation x^* , for each agent in group i at most $(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} c_{i,j}(\mathbf{k})$ dis-utility has been removed from the bundle of an other group and has been allocated to group i .

Thus $\min_{i' \in [d]} \text{CGAP}_{i,i'}(A) \geq \min_{i' \in [d]} \text{CGAP}_{i,i'}(x^*) - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} c_{i,j}(\mathbf{k}) = \|c_i(\mathbf{k})\|_2 \cdot \alpha^* - 2(d(d-1) + t(\theta + n + n_d - d - 1)) \max_{j \in [t]} c_{i,j}(\mathbf{k})$. $\square$

Proof of Lemma 9. It is easy to verify that $\alpha \|c_i\|_1 \leq \|c_i(\mathbf{k})\|_1 \leq \beta \|c_i\|_1$. Thus there exist $\alpha \leq \gamma_{i'} \leq \beta$ and $\alpha \leq \gamma_i \leq \beta$ such that $\|c_i(\mathbf{k})\|_1 = \gamma_i \cdot \|c_i\|_1$ and $\|c_{i'}(\mathbf{k})\|_1 = \gamma_{i'} \cdot \|c_{i'}\|_1$.

$$\begin{aligned}
D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_{i'}^{(1)}(\mathbf{k})) &= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{\tilde{c}_{i,j}^{(1)}(\mathbf{k})}{\tilde{c}_{i',j}^{(1)}(\mathbf{k})} \right) \\
&= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{c_{i,j}(\mathbf{k}) / \|c_i(\mathbf{k})\|_1}{c_{i',j}(\mathbf{k}) / \|c_{i'}(\mathbf{k})\|_1} \right) \\
&= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{c_{i,j}(\mathbf{k}) / \|c_i\|_1}{c_{i',j}(\mathbf{k}) / \|c_{i'}\|_1} \cdot \frac{\gamma_{i'}}{\gamma_i} \right) \\
&= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)}(\mathbf{k}) \ln \left(\frac{c_{i,j}(\mathbf{k}) / \|c_i\|_1}{c_{i',j}(\mathbf{k}) / \|c_{i'}\|_1} \right) \\
&= \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \sum_{j \in \mathcal{M}} \frac{c_{i,j}(\mathbf{k})}{\|c_i(\mathbf{k})\|_1} \ln \left(\frac{c_{i,j}(\mathbf{k}) / \|c_i\|_1}{c_{i',j}(\mathbf{k}) / \|c_{i'}\|_1} \right) \\
&= \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{1}{\gamma_i} \sum_{z \in [t]} k_z \cdot \frac{c_{i,z}}{\|c_i\|_1} \ln \left(\frac{c_{i,z} / \|c_i\|_1}{c_{i',z} / \|c_{i'}\|_1} \right) \\
&= \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{1}{\gamma_i} \sum_{z \in [t]} k_z \cdot \tilde{c}_{i,z}^{(1)} \ln \left(\frac{\tilde{c}_{i,z}^{(1)}}{\tilde{c}_{i',z}^{(1)}} \right) \\
&= \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{1}{\gamma_i} \left(\sum_{z \in [t]} k_z \cdot \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{i,z}^{(1)} \right) - \sum_{z \in [t]} k_z \cdot \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{i',z}^{(1)} \right) \right) \\
&\geq \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{1}{\gamma_i} \left(\beta \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{i,z}^{(1)} \right) - \alpha \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{i',z}^{(1)} \right) \right) \quad (0 < \tilde{c}_{i,z}^{(1)} \leq 1) \\
&= \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{\alpha}{\gamma_i} \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\frac{\tilde{c}_{i,z}^{(1)}}{\tilde{c}_{i',z}^{(1)}} \right) + \frac{\beta - \alpha}{\gamma_i} \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{i,z}^{(1)} \right) \\
&\geq \ln \left(\frac{\gamma_{i'}}{\gamma_i} \right) + \frac{\alpha}{\gamma_i} \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\frac{\tilde{c}_{i,z}^{(1)}}{\tilde{c}_{i',z}^{(1)}} \right) + \frac{\beta - \alpha}{\gamma_i} \sum_{z \in [t]} \tilde{c}_{i,z}^{(1)} \ln \left(\tilde{c}_{min}^{(1)} \right) \\
&\geq \frac{\alpha}{\beta} D_{KL}(\tilde{c}_i^{(1)} \parallel \tilde{c}_{i'}^{(1)}) - \left(\ln \left(\frac{\beta}{\alpha} \right) + \frac{\beta - \alpha}{\alpha} \ln \left(\tilde{c}_{min}^{(1)} \right)^{-1} \right).
\end{aligned}$$

$\square$

Proof of Lemma 10.

$$\begin{aligned}
x \left(\frac{x}{x+a} b - \ln \frac{x+a}{x} \right) - c \ln(x+a) &= x \left(\frac{x}{x+a} b - \ln(1 + a/x) \right) - c \ln(x+a) \\
&\geq x \left(\frac{x}{x+a} b - \frac{a}{x} \right) - c \ln(x+a) \quad (\ln(1+u) \leq u) \\
&\geq x \left(\frac{x-a}{x} b - \frac{a}{x} \right) - c \ln(x+a) \quad \left(\frac{x-a}{x} \leq \frac{x}{x+a} \right) \\
&= xb - c \ln(x+a) - a(b+1)
\end{aligned}$$

Now for function $f(x) = xb - c \ln(x + a)$, notice that $f''(x) = \frac{c}{(a+x)^2} \geq 0$. Thus, $f(x)$ is convex, which implies that it always lies above its tangent lines. We pick the tangent for $x = \frac{2c}{b} - a$ since it simplifies our bounds. Combining the above we have that:

$$\begin{aligned}
x \left(\frac{x}{x+a} b - \ln \frac{x+a}{x} \right) - c \ln(x+a) &\geq xb - c \ln(x+a) - a(b+1) \\
&= f(x) - a(b+1) \\
&\geq f\left(\frac{2c}{b} - a\right) + f'\left(\frac{2c}{b} - a\right) \cdot \left(x - \left(\frac{2c}{b} - a\right)\right) - a(b+1) \\
&= 2c - ab - c \ln\left(\frac{2c}{b}\right) + \frac{b}{2}x - c + \frac{ab}{2} - ab - a \\
&= \frac{b}{2}x - \left(a\left(\frac{3}{2}b + 1\right) + c\left(\ln\left(\frac{2c}{b}\right) - 1\right)\right).
\end{aligned}$$

□

Proof of Theorem 5. Let $\mu = 2 \frac{n + (\frac{5}{2}n + \lambda - 1) \ln(\tilde{c}_{min}^{(1)})^{-1} + \lambda \left(\ln\left(\frac{2\lambda}{\min_{\hat{i}, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_i^{(1)})} - 1 \right) \right)}{\min_{\hat{i}, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_i^{(1)})}$. First, notice that, without loss of generality, we can assume that $k_z \in [\mu, \mu + n - 1]$. If that is not the case, and $k_z > \mu + n - 1$, we can simply allocate one copy of item z to each agent and continue this process until the condition is satisfied; since all agents receive the exact same amount of items, envy is not affected by this process. Also, since we allocate copies in multiples of n and $n \equiv 0 \pmod{g}$, then the remaining number of copies is also divisible by g (so, the condition $k_z \equiv 0 \pmod{g}$ also remains true).

From Theorem 4 we know that if $\tilde{c}_{max}^{(1)}(\mathbf{k}) \leq \frac{\min_{\hat{i}, \tilde{i} \in [d]} D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) || \tilde{c}_i^{(1)}(\mathbf{k}))}{\lambda \ln \tilde{c}_{min}^{(1)}(\mathbf{k})^{-1}}$, then an envy-free allocation exists. Notice that $\tilde{c}_{max}^{(1)}(\mathbf{k}) \leq \frac{\tilde{c}_{max}^{(1)}}{\mu} \leq \frac{1}{\mu}$. Thus, we conclude that if

$$\mu \cdot \min_{i \in [d], i' \in [d]} D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) || \tilde{c}_{i'}^{(1)}(\mathbf{k})) \geq \lambda \ln \tilde{c}_{min}^{(1)}(\mathbf{k})^{-1}$$

an envy-free allocation exists. We also have that for any $\hat{i}, \tilde{i} \in [d]$:

$$\begin{aligned}
0 &= \mu \frac{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})}{2} - \left(n + \left(\frac{5}{2}n + \lambda - 1 \right) \ln(\tilde{c}_{min}^{(1)})^{-1} + \lambda \left(\ln\left(\frac{2\lambda}{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})} - 1 \right) \right) \right) \\
&= \mu \frac{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})}{2} - n \left(\frac{3}{2} \ln(\tilde{c}_{min}^{(1)})^{-1} + 1 \right) - (n-1) \ln(\tilde{c}_{min}^{(1)})^{-1} - \lambda \ln(\tilde{c}_{min}^{(1)})^{-1} \\
&\quad - \lambda \left(\ln\left(\frac{2\lambda}{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})} - 1 \right) \right) \\
&\leq \mu \frac{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})}{2} - (n-1) \left(\frac{3D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})}{2} + 1 \right) - (n-1) \ln(\tilde{c}_{min}^{(1)})^{-1} \\
&\quad - \lambda \ln(\tilde{c}_{min}^{(1)})^{-1} - \lambda \left(\ln\left(\frac{2\lambda}{D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)})} - 1 \right) \right) \\
&\quad (D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)}) \leq \ln(\tilde{c}_{min}^{(1)})^{-1}) \\
&\leq \mu \left(\frac{\mu}{\mu + n - 1} D_{KL}(\tilde{c}_i^{(1)} || \tilde{c}_{\tilde{i}}^{(1)}) - \ln\left(\frac{\mu + n - 1}{\mu}\right) \right) - \lambda \ln(\mu + n - 1) - \lambda \ln(\tilde{c}_{min}^{(1)})^{-1} \\
&\quad - (n-1) \ln(\tilde{c}_{min}^{(1)})^{-1} \tag{Lemma 10}
\end{aligned}$$

$$\begin{aligned}
&= \mu \left(\frac{\mu}{\mu+n-1} D_{KL}(\tilde{c}_i^{(1)} \parallel \tilde{c}_i^{(1)}) - \ln \left(\frac{\mu+n-1}{\mu} \right) - \frac{n-1}{\mu} \ln \left(\tilde{c}_{min}^{(1)} \right)^{-1} \right) - \lambda \ln \left(\frac{\tilde{c}_{min}^{(1)}}{\mu+n-1} \right)^{-1} \\
&\leq \mu D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_i^{(1)}(\mathbf{k})) - \lambda \ln \left(\frac{\tilde{c}_{min}^{(1)}}{\mu+n-1} \right)^{-1} \quad (\text{Lemma 9}) \\
&\leq \mu D_{KL}(\tilde{c}_i^{(1)}(\mathbf{k}) \parallel \tilde{c}_i^{(1)}(\mathbf{k})) - \lambda \ln \left(\tilde{c}_{min}^{(1)}(\mathbf{k}) \right)^{-1} \quad (\tilde{c}_{min}^{(1)}(\mathbf{k}) \geq \frac{\tilde{c}_{min}^{(1)}}{\mu+n-1})
\end{aligned}$$

Combining the above, we get that an envy-free allocation exists. $\square$

C Proofs missing from Section 4.2

Proof of Lemma 12. Consider the function $\bar{f}_i(x) = \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} \tilde{u}_i^{(2)}(y) dy$.

$$\begin{aligned}
|\tilde{u}_i^{(2)}(x) - \bar{f}_i(x)| &= \left| \tilde{u}_i^{(2)}(x) - \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} \tilde{u}_i^{(2)}(y) dy \right| \\
&= \left| \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} \tilde{u}_i^{(2)}(x) dy - \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} \tilde{u}_i^{(2)}(y) dy \right| \\
&= \left| \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} \tilde{u}_i^{(2)}(x) - \tilde{u}_i^{(2)}(y) dy \right| \\
&\leq \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} |\tilde{u}_i^{(2)}(x) - \tilde{u}_i^{(2)}(y)| dy \\
&\leq \frac{1}{\epsilon} \int_{\lfloor x/\epsilon \rfloor \epsilon}^{\lfloor x/\epsilon \rfloor \epsilon + 1} k|x-y| dy \\
&= \frac{k}{\epsilon} \left(\int_{\lfloor x/\epsilon \rfloor \epsilon}^x (x-y) dy + \int_x^{\lfloor x/\epsilon \rfloor \epsilon + 1} (y-x) dy \right) \\
&= \frac{k}{\epsilon} \left(x(x - \lfloor x/\epsilon \rfloor \epsilon) - x((\lfloor x/\epsilon \rfloor + 1)\epsilon - x) - \frac{1}{2} [x^2 - (\lfloor x/\epsilon \rfloor \epsilon)^2] + \frac{1}{2} [((\lfloor x/\epsilon \rfloor + 1)\epsilon)^2 - x^2] \right) \\
&= \frac{k}{\epsilon} \left(x(2x - (2\lfloor x/\epsilon \rfloor + 1)\epsilon) - x^2 + \frac{1}{2} ((\lfloor x/\epsilon \rfloor \epsilon)^2 + ((\lfloor x/\epsilon \rfloor + 1)\epsilon)^2) \right) \\
&= \frac{k}{\epsilon} \left(x^2 - x(2\lfloor x/\epsilon \rfloor + 1)\epsilon + (\lfloor x/\epsilon \rfloor \epsilon)^2 + \lfloor x/\epsilon \rfloor \epsilon^2 + \frac{1}{2}\epsilon^2 \right) \\
&= \frac{k}{\epsilon} \left(\left(x - \lfloor x/\epsilon \rfloor \epsilon - \frac{\epsilon}{2} \right)^2 + \frac{\epsilon^2}{4} \right) \\
&\leq \frac{k\epsilon}{2}
\end{aligned}$$

$$\sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x) - \bar{f}_i(x))^2 dx} \leq \sqrt{\int_0^1 \left(\frac{k\epsilon}{2} \right)^2 dx} = \left(\frac{k\epsilon}{2} \right)$$

We have the following:

$$\begin{aligned}
\sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x))^2 dx} &= \sqrt{\int_0^1 (\bar{f}_i(x) + (\tilde{u}_i^{(2)}(x) - \bar{f}_i(x)))^2 dx} \\
&\leq \sqrt{\int_0^1 (\bar{f}_i(x))^2 dx} + \sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x) - \bar{f}_i(x))^2 dx} \quad (\text{Minkowski's Inequality})
\end{aligned}$$

$$\begin{aligned}
&\leq \sqrt{\int_0^1 (\bar{f}_i(x))^2 dx} + \sqrt{\int_0^1 \left(\frac{k\epsilon}{2}\right)^2 dx} \\
&= \sqrt{\int_0^1 (\bar{f}_i(x))^2 dx} + \frac{k\epsilon}{2}
\end{aligned}$$

Thus $\sqrt{\int_0^1 (\bar{f}_i(x))^2 dx} \geq \sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x))^2 dx} - \frac{k\epsilon}{2} = 1 - \frac{k\epsilon}{2}$. Similarly we can get that:

$$\begin{aligned}
\sqrt{\int_0^1 (\bar{f}_i(x))^2 dx} &= \sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x) + (\bar{f}_i(x) - \tilde{u}_i^{(2)}(x)))^2 dx} \\
&\leq \sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x))^2 dx} + \sqrt{\int_0^1 (\bar{f}_i(x) - \tilde{u}_i^{(2)}(x))^2 dx} \quad (\text{Minkowski's Inequality}) \\
&\leq \sqrt{\int_0^1 (\tilde{u}_i^{(2)}(x))^2 dx} + \sqrt{\int_0^1 \left(\frac{k\epsilon}{2}\right)^2 dx} \\
&= 1 + \frac{k\epsilon}{2}
\end{aligned}$$

$$\begin{aligned}
&\int_0^1 (\tilde{u}_i^{(2)}(x) - \tilde{u}_{i'}^{(2)}(x))^2 dx - \sum_{j=0}^{1/\epsilon-1} \left(\frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}} - \frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} \right)^2 \\
&= \int_0^1 (\tilde{u}_i^{(2)}(x))^2 + (\tilde{u}_{i'}^{(2)}(x))^2 - 2\tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \\
&\quad - \sum_{j=0}^{1/\epsilon-1} \left(\frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}} - \frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} \right)^2 \\
&= 2 - 2 \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx - \sum_{j=0}^{1/\epsilon-1} \left(\frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2}} - \frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} \right)^2 \\
&= 2 - 2 \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx - \sum_{j=0}^{1/\epsilon-1} \left(\frac{\left(\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx \right)^2}{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2} + \frac{\left(\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2} \right. \\
&\quad \left. - 2 \frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx \cdot \int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2} \cdot \sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} \right) \\
&= 2 \left(\sum_{j=0}^{1/\epsilon-1} \frac{\int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_i^{(2)}(x) dx \cdot \int_{\epsilon j}^{\epsilon(1+j)} \tilde{u}_{i'}^{(2)}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2} \cdot \sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\sum_{j=0}^{1/\epsilon-1} \frac{\epsilon^2 \bar{f}_i(j) \cdot \bar{f}_{i'}(j)}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2} \cdot \sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right)
\end{aligned}$$

$$\begin{aligned}
&= 2 \left(\frac{\epsilon \sum_{j=0}^{1/\epsilon-1} \int_{\epsilon j}^{\epsilon(1+j)} \bar{f}_i(\epsilon j) \cdot \bar{f}_{i'}(\epsilon j) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2 \cdot \sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\frac{\epsilon \sum_{j=0}^{1/\epsilon-1} \int_{\epsilon j}^{\epsilon(1+j)} \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2 \cdot \sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&\quad (\tilde{u}_i^{(2)}(\epsilon j) = \tilde{u}_i^{(2)}(x) \text{ for } x \in [\epsilon j, \epsilon(j+1)]) \\
&= 2 \left(\frac{\epsilon \int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_i^{(2)}(x) dx \right)^2 \cdot \sum_{j'=0}^{1/\epsilon-1} \left(\int_{\epsilon j'}^{\epsilon(j'+1)} \tilde{u}_{i'}^{(2)}(x) dx \right)^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\frac{\epsilon \int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} (\epsilon \bar{f}_i(\epsilon j'))^2 \cdot \sum_{j'=0}^{1/\epsilon-1} (\epsilon \bar{f}_{i'}(\epsilon j'))^2}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\frac{\epsilon \int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\epsilon \sum_{j'=0}^{1/\epsilon-1} \int_{\epsilon j'}^{\epsilon(j'+1)} (\bar{f}_i(\epsilon j'))^2 dx \cdot \epsilon \sum_{j'=0}^{1/\epsilon-1} \int_{\epsilon j'}^{\epsilon(j'+1)} (\bar{f}_{i'}(\epsilon j'))^2 dx}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\frac{\int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\sum_{j'=0}^{1/\epsilon-1} \int_{\epsilon j'}^{\epsilon(j'+1)} (\bar{f}_i(x))^2 dx \cdot \sum_{j'=0}^{1/\epsilon-1} \int_{\epsilon j'}^{\epsilon(j'+1)} (\bar{f}_{i'}(x))^2 dx}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= 2 \left(\frac{\int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\sqrt{\int_0^1 (\bar{f}_i(x))^2 dx \cdot \int_0^1 (\bar{f}_{i'}(x))^2 dx}} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&\leq \left(\frac{\int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx}{\left(1 - \frac{k\epsilon}{2}\right)} - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) \\
&= \left(\int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx - \int_0^1 \tilde{u}_i^{(2)}(x) \cdot \tilde{u}_{i'}^{(2)}(x) dx \right) + \frac{\frac{k\epsilon}{2}}{\left(1 - \frac{k\epsilon}{2}\right)} \int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx \\
&= \left(\int_0^1 \bar{f}_i(x) \cdot \left(\bar{f}_{i'}(x) - \tilde{u}_{i'}^{(2)}(x) \right) dx + \int_0^1 \tilde{u}_{i'}^{(2)}(x) \cdot \left(\bar{f}_i(x) - \tilde{u}_i^{(2)}(x) \right) dx \right) + \frac{\frac{k\epsilon}{2}}{\left(1 - \frac{k\epsilon}{2}\right)} \int_0^1 \bar{f}_i(x) \cdot \bar{f}_{i'}(x) dx \\
&\leq \sqrt{\int_0^1 (\bar{f}_i(x))^2 dx \cdot \int_0^1 \left(\bar{f}_{i'}(x) - \tilde{u}_{i'}^{(2)}(x) \right)^2 dx} + \sqrt{\int_0^1 \left(\tilde{u}_{i'}^{(2)}(x) \right)^2 dx \cdot \int_0^1 \left(\bar{f}_i(x) - \tilde{u}_i^{(2)}(x) \right)^2 dx} \\
&\quad + \frac{k\epsilon}{2 - k\epsilon} \sqrt{\int_0^1 (\bar{f}_i(x))^2 dx \cdot \int_0^1 (\bar{f}_{i'}(x))^2 dx} \quad (\text{Cauchy-Schwartz}) \\
&\leq \left(1 + \frac{k\epsilon}{2} \right) \frac{k\epsilon}{2} + \frac{k\epsilon}{2} + \frac{k\epsilon}{2 - k\epsilon} \left(1 + \frac{k\epsilon}{2} \right)^2 \\
&\leq 3.5 k\epsilon. \quad (\text{Assuming } k\epsilon \leq 1)
\end{aligned}$$

□

D Proofs missing from Section 4.3

Proof of Lemma 15. From the variational characterization of chi-squared divergence, we know that $D_{\chi^2}(\tilde{v}_i^{(1)} || S) \geq \langle \tilde{v}_i^{(1)}, g \rangle - \frac{\langle S, g^2 \rangle}{4} - 1$, where g is any vector in $\mathbb{R}^m$ and $g^2 = [g_1^2, g_2^2, \dots, g_m^2]$. By choosing $g = 2m \cdot \tilde{v}_i^{(1)}$, we

have that $D_{\chi^2}(\tilde{v}_i^{(1)}||S) \geq \langle \tilde{v}_i^{(1)}, 2m \cdot \tilde{v}_i^{(1)} \rangle - \frac{\langle S, (2m \cdot \tilde{v}_i^{(1)})^2 \rangle}{4} - 1$.

From Hoeffding's bound, and using the fact that $\mathbb{E}[v_{i,j}^2] = \frac{1}{3}$ and $\mathbb{E}[v_{i,j}^3] = \frac{1}{4}$, we have the following bounds, for all $i \in \mathcal{N}$: (i) $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i,j'} \geq \frac{m}{2} + \sqrt{m \log n}\right] \leq 1/n^2$, (ii) $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i,j'}^2 \leq \frac{m}{3} - \sqrt{m \log n}\right] \leq 1/n^2$, (iii) $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i,j'}^2 \geq \frac{m}{3} + \sqrt{m \log n}\right] \leq 1/n^2$, as well as (iv) $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i,j'}^3 \geq \frac{m}{4} + \sqrt{m \log n}\right] \leq 1/n^2$. Also from Hoeffding's bound we have that for any $i' \in \mathcal{N} \setminus \{i\}$: $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i',j'} \leq m/2 - \sqrt{m \log n}\right] \leq 1/n^2$ and $\mathbb{P}\left[\sum_{j' \in \mathcal{M}} v_{i',j'} \cdot v_{i,j'}^2 \geq m/6 + \sqrt{m \log n}\right] \leq 1/n^2$.

None of these events occurs with high probability; conditioning on this, we have:

$$\begin{aligned}
D_{\chi^2}(v_i||S) &\geq \langle \tilde{v}_i^{(1)}, 2m \cdot \tilde{v}_i^{(1)} \rangle - \frac{\langle S, (2m \cdot \tilde{v}_i^{(1)})^2 \rangle}{4} - 1 \\
&= \sum_{j \in \mathcal{M}} 2m \cdot (\tilde{v}_{i,j}^{(1)})^2 - \sum_{j \in \mathcal{M}} m^2 \cdot (S_j(\tilde{v}_{i,j}^{(1)}))^2 - 1 \\
&= 2m \sum_{j \in \mathcal{M}} \frac{v_{i,j}^2}{\left(\sum_{j' \in \mathcal{M}} v_{i,j'}\right)^2} - \frac{m^2}{n} \sum_{j \in \mathcal{M}} \sum_{i' \in \mathcal{N}} \tilde{v}_{i',j}^{(1)} \cdot (\tilde{v}_{i,j}^{(1)})^2 - 1 \\
&= 2m \sum_{j \in \mathcal{M}} \frac{v_{i,j}^2}{\left(\sum_{j' \in \mathcal{M}} v_{i,j'}\right)^2} - \frac{m^2}{n} \sum_{j \in \mathcal{M}} \sum_{i' \in \mathcal{N} \setminus \{i\}} \frac{v_{i',j}}{\sum_{j' \in \mathcal{M}} v_{i',j'}} \cdot \frac{v_{i,j}^2}{\left(\sum_{j' \in \mathcal{M}} v_{i,j'}\right)^2} - \frac{m^2}{n} \sum_{j \in \mathcal{M}} \frac{v_{i,j}^3}{\left(\sum_{j' \in \mathcal{M}} v_{i,j'}\right)^3} - 1 \\
&\geq \frac{2m}{(m/2 + \sqrt{m \log n})^2} \sum_{j \in \mathcal{M}} v_{i,j}^2 - \frac{m^2}{n(m/2 - \sqrt{m \log n})^3} \sum_{j \in \mathcal{M}} \sum_{i' \in \mathcal{N} \setminus \{i\}} v_{i',j} v_{i,j}^2 - \frac{m^2}{n(m/2 - \sqrt{m \log n})^2} \sum_{j \in \mathcal{M}} v_{i,j}^3 - 1 \\
&\geq \frac{2m(m/3 - \sqrt{m \log n})}{(m/2 + \sqrt{m \log n})^2} - \frac{m^2(n-1)(m/6 + \sqrt{m \log n})}{n(m/2 - \sqrt{m \log n})^3} - \frac{m^2(m/4 + \sqrt{m \log n})}{n(m/2 - \sqrt{m \log n})^3} - 1 \\
&= \frac{8}{3} \frac{1 - 3\sqrt{\frac{\log n}{m}}}{\left(1 + 2\sqrt{\frac{\log n}{m}}\right)^2} - \frac{4(n-1)}{3n} \frac{1 + 6\sqrt{\frac{\log n}{m}}}{\left(1 - 2\sqrt{\frac{\log n}{m}}\right)^3} - \frac{2}{n} \frac{1 + 4\sqrt{\frac{\log n}{m}}}{\left(1 - 2\sqrt{\frac{\log n}{m}}\right)^3} - 1.
\end{aligned}$$

Let $\delta = \sqrt{\frac{\log n}{m}}$. Then, we have that

$$\begin{aligned}
D_{\chi^2}(v_i||S) &\geq \frac{8}{3} \frac{1 - 3\delta}{(1 + 2\delta)^2} - \frac{4(n-1)}{3n} \frac{1 + 6\delta}{(1 - 2\delta)^3} - \frac{2}{n} \frac{1 + 4\delta}{(1 - 2\delta)^3} - 1 \\
&\geq \frac{8}{3} \frac{1 - 3\delta}{(1 + 2\delta)^2} - \frac{1 + 6\delta}{(1 - 2\delta)^3} \left(\frac{4(n-1)}{3n} + \frac{2}{n} \right) - 1 \\
&= \frac{8}{3} \frac{1 - 3\delta}{(1 + 2\delta)^2} - \frac{1 + 6\delta}{(1 - 2\delta)^3} \left(\frac{4n+2}{3n} \right) - 1
\end{aligned}$$

For $m \geq (11z - 11)^2 \log(n)$ we have that $\frac{1-3\delta}{(1+2\delta)^2} \geq \frac{z-1}{z}$ ($z \geq 2$), and for $m \geq (12z + 6)^2 \log(n)$ we have that $\frac{1+6\delta}{(1-2\delta)^3} \leq \frac{z+1}{z}$. Therefore, for $m \geq \max\{(11z - 11)^2, (12z + 6)^2\} \log(n)$, we have that

$$\begin{aligned}
D_{\chi^2}(v_i||S) &\geq \frac{8}{3} \frac{z-1}{z} - \frac{z+1}{z} \left(\frac{4n+2}{3n} \right) - 1 \\
&= \frac{1}{z} \frac{8(z-1)n - 4n(z+1) - 2(z+1) - 3nz}{3n} \\
&= \frac{1}{z} \frac{n(7z-12) - 2(z+1)}{3n}.
\end{aligned}$$

By picking $z = 2$, we have that, for $m \geq 900 \log(n)$, $D_{\chi^2}(v_i||S) \geq \frac{n-3}{3n}$. $\square$

Proof of Lemma 16.

$$\begin{aligned}
\frac{c_i(x_i)}{c_i(\mathcal{M})} &= \frac{1}{c_i(\mathcal{M})} \sum_{j \in \mathcal{M}} c_{i,j} \frac{1/\tilde{c}_{i,j}^{(1)}}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}} \\
&= \sum_{j \in \mathcal{M}} \frac{1}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}} \\
&= \frac{1}{n} \sum_{j \in \mathcal{M}} \frac{n}{\sum_{i' \in \mathcal{N}} 1/\tilde{c}_{i',j}^{(1)}} \\
&= \frac{1}{n} \sum_{j \in \mathcal{M}} H_j \quad (H_j = \frac{n}{\sum_{i' \in \mathcal{N}} \frac{1}{\tilde{c}_{i',j}^{(1)}}}) \\
&= \frac{1}{n} \left(1 - 1 + \sum_{j \in \mathcal{M}} (2H_j - H_j) \right) \\
&= \frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \tilde{c}_{\ell,j}^{(1)} + \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} 2H_j - \frac{1}{n} \sum_{j \in \mathcal{M}} H_j^2 \frac{n}{H_j} \right) \quad (\sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} = 1) \\
&= \frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \tilde{c}_{\ell,j}^{(1)} + \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} 2H_j - \frac{1}{n} \sum_{j \in \mathcal{M}} H_j^2 \sum_{\ell \in \mathcal{N}} \frac{1}{\tilde{c}_{\ell,j}^{(1)}} \right) \quad (H_j = \frac{n}{\sum_{\ell \in \mathcal{N}} \frac{1}{\tilde{c}_{\ell,j}^{(1)}}}) \\
&= \frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)})^2 - 2H_j \tilde{c}_{\ell,j}^{(1)} + H_j^2}{\tilde{c}_{\ell,j}^{(1)}} \right) \\
&= \frac{1}{n} \left(1 - \frac{1}{n} \sum_{\ell \in \mathcal{N}} \sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{\ell,j}^{(1)} - H_j)^2}{\tilde{c}_{\ell,j}^{(1)}} \right).
\end{aligned}$$

□

Proof of Lemma 18.

$$\begin{aligned}
\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} &= \sum_{j \in \mathcal{M}} \frac{H_j^2}{\tilde{c}_{i,j}^{(1)}} - 2H_j + \tilde{c}_{i,j}^{(1)} \\
&= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} \left(\left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}} \right)^2 - 2\frac{H_j}{\tilde{c}_{i,j}^{(1)}} + 1 \right)
\end{aligned}$$

Now consider the function $f(t) = t^2 - 2t + 1$ and its Legendre-Fenchel transformation $f^*(s)$:

$$f^*(s) = \sup_{t \in \mathbb{R}} st - f(t) = \sup_{t \in \mathbb{R}} st - t^2 + 2t - 1 = \frac{s^2}{4} + s,$$

where the last equality is derived by taking the first and second derivatives and observing that $st - t^2 + 2t - 1$ is maximized for $t = s/2 + 1$. Thus $f(t) \geq st - f^*(s) = s(t - 1) - \frac{s^2}{4}$. Setting $t = \frac{H_j}{\tilde{c}_{i,j}^{(1)}}$ and $s = g_j$ we have

$$f\left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}}\right) = \left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}}\right)^2 - 2\frac{H_j}{\tilde{c}_{i,j}^{(1)}} + 1 \geq g_j \left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}} - 1\right) - \frac{1}{4}g_j^2.$$

Combining all of the above we have that:

$$\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} = \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} \left(\left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}} \right)^2 - 2\frac{H_j}{\tilde{c}_{i,j}^{(1)}} + 1 \right)$$

$$\begin{aligned}
&= \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} f\left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}}\right) \\
&\geq \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} \left(g_j \left(\frac{H_j}{\tilde{c}_{i,j}^{(1)}} - 1 \right) - \frac{1}{4} g_j^2 \right) \\
&= \sum_{j \in \mathcal{M}} g_j (H_j - \tilde{c}_{i,j}^{(1)}) + \frac{1}{4} \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} g_j^2 \\
&= \langle H - \tilde{c}_i^{(1)}, g \rangle - \frac{\langle \tilde{c}_i^{(1)}, g^2 \rangle}{4}.
\end{aligned}$$

Interestingly, this lower bound on $\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}}$ is tight; this can be verified that by choosing $g_j = 2 \frac{H_j}{\tilde{c}_{i,j}^{(1)}} - 2$. $\square$

Proof of Lemma 19. Using standard Chernoff-Hoeffding bounds we have $\mathbb{P}\left[\sum_{j \in \mathcal{M}} c_{i,j} \leq m/2 - \sqrt{m \log n}\right] \leq 1/n^2$, for all $i \in \mathcal{N}$. Therefore, with high probability, $\sum_{j \in \mathcal{M} \setminus \{j'\}} c_{i,j} \geq m/2 - 1 - \sqrt{m \log n}$ for all $j' \in \mathcal{M}$. Now consider the random variable $G_j = \left(\prod_{i \in [n]} c_{i,j}\right)^{1/n}$. It holds that $0 \leq G_j \leq 1$, and $\mathbb{E}[G_j] = \left(\frac{n}{n+1}\right)^n$. Therefore, by standard Chernoff-Hoeffding bounds, we have that $\mathbb{P}\left[\sum_{j \in \mathcal{M}} G_j \geq m \left(\frac{n}{n+1}\right)^n + \sqrt{m \log n}\right] \leq 1/n^2$ for all $i \in \mathcal{N}$. Union bounding, with probability $1 - O(1/n)$, all of the above inequalities hold.

Lemma 18 gives us that $\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} \geq \langle H - \tilde{c}_i^{(1)}, g \rangle - \frac{\langle \tilde{c}_i^{(1)}, g^2 \rangle}{4}$. Choosing $g_{i,j} = -\frac{1}{2}$ we have:

$$\begin{aligned}
\sum_{j \in \mathcal{M}} \frac{(\tilde{c}_{i,j}^{(1)} - H_j)^2}{\tilde{c}_{i,j}^{(1)}} &\geq \langle H - \tilde{c}_i^{(1)}, g \rangle - \frac{\langle \tilde{c}_i^{(1)}, g^2 \rangle}{4} \\
&= \sum_{j \in \mathcal{M}} \frac{1}{2} (\tilde{c}_{i,j}^{(1)} - H_j) - \frac{1}{4} \sum_{j \in \mathcal{M}} \frac{1}{4} \tilde{c}_{i,j}^{(1)} \\
&= \frac{7}{16} \sum_{j \in \mathcal{M}} \tilde{c}_{i,j}^{(1)} - \frac{1}{2} \sum_{j \in \mathcal{M}} H_j \\
&= \frac{7}{16} - \frac{1}{2} \sum_{j \in \mathcal{M}} H_j \\
&\geq \frac{7}{16} - \frac{1}{2} \sum_{j \in \mathcal{M}} \left(\prod_{i \in [n]} \tilde{c}_{i,j}^{(1)} \right)^{1/n} && \text{(HM-GM inequality)} \\
&= \frac{7}{16} - \frac{1}{2} \sum_{j \in \mathcal{M}} \left(\prod_{i \in [n]} \frac{c_{i,j}}{\sum_{j' \in \mathcal{M}} c_{i,j'}} \right)^{1/n} \\
&\geq \frac{7}{16} - \frac{1}{2} \sum_{j \in \mathcal{M}} \left(\prod_{i \in [n]} \frac{c_{i,j}}{\sum_{j' \in \mathcal{M} \setminus \{j\}} c_{i,j'}} \right)^{1/n} && (0 \leq c_{i,j} \leq 1) \\
&\geq \frac{7}{16} - \frac{1}{2} \sum_{j \in \mathcal{M}} \left(\prod_{i \in [n]} \frac{c_{i,j}}{\frac{m}{2} - 1 - \sqrt{m \log n}} \right)^{1/n} && \text{(Hoeffding high-probability bound)} \\
&= \frac{7}{16} - \frac{1}{m - 2 - 2\sqrt{m \log n}} \sum_{j \in \mathcal{M}} \left(\prod_{i \in [n]} c_{i,j} \right)^{1/n}
\end{aligned}$$

$$\begin{aligned}
&\geq \frac{7}{16} - \frac{m \left(\frac{n}{n+1}\right)^n + \sqrt{m \log n}}{m - 2 - 2\sqrt{m \log n}} && \text{(Hoeffding high-probability bound)} \\
&= \frac{7}{16} - \frac{\left(\frac{n}{n+1}\right)^n + \sqrt{\frac{\log n}{m}}}{1 - \frac{2}{m} - 2\sqrt{\frac{\log n}{m}}} \\
&\geq \frac{7}{16} - \frac{\left(\frac{n}{n+1}\right)^n + \sqrt{\frac{\log n}{m}}}{1 - 3\sqrt{\frac{\log n}{m}}} \\
&\geq \frac{7}{16} - \frac{\frac{4}{9} + \sqrt{\frac{\log n}{m}}}{1 - 3\sqrt{\frac{\log n}{m}}} && \left(\left(\frac{n}{n+1}\right)^n \text{ is decreasing}\right) \\
&\geq \frac{5}{16}.
\end{aligned}$$

□

E Allocating Goods to Similar Agents

So far, we have considered the situation where agents' valuation vectors are dissimilar and proved that fair allocations exist. It is well understood that when agents are identical, EFX allocations exist. As a proof of concept, in this section we prove that when agents are similar — but not identical — EFX allocations of goods exist, complementing our previous results.

We consider the fairness notion of *tEFX*, or “transfer EFX”. An integral allocation A is *tEFX* if, for all $i, i' \in \mathcal{N}$, for all goods $j \in A_{i'}$, such that $v_i(A_i) + v_{i,j} \geq v_i(A_{i'}) - v_{i,j}$. Note that this notion is stronger than (i.e., it implies) *EF2x*, which demands that the removal of any two distinct goods eliminates envy, and was recently shown to exist for four additive agents [AGS25]. Here, we prove that *tEFX* allocations exist for any number of agents that are “pairwise similar.” Formally, we require a bound on the TV distance between $\tilde{v}_i^{(1)}$ and $\tilde{v}_{i'}^{(1)}$ for any pair of agents $i, i' \in \mathcal{N}$.

We use the following greedy algorithm. Algorithm 1 allocates the item that is the least desirable to agent 1 to the agent with the smallest utility from the point of view of agent 1; note that if $v_1(S) \geq v_1(S')$, then $\tilde{v}_1^{(1)}(S) \geq \tilde{v}_1^{(1)}(S')$, for any two bundles S and S' . This algorithm produces an EFX allocation for identical agents. Here we prove that it also produces a *tEFX* allocation for non-identical, but “similar,” agents.

Algorithm 1 Greedy Allocation

Initialize $A = (A_1, \dots, A_n)$, with $A_i = \emptyset$.
Rename the items by decreasing order in terms of $v_{1,j}$ (i.e., $v_{1,1} \geq v_{1,2} \geq \dots$).
for each item j in the sorted order **do**
 $i \leftarrow \arg \min_{i' \in \mathcal{N}} v_1(A_{i'})$
 $A_i \leftarrow A_i \cup \{j\}$
end for

Theorem 11. *Let $D^* = \max_{i,i' \in \mathcal{N}} D_{TV}(\tilde{v}_i^{(1)} \parallel \tilde{v}_{i'}^{(1)})$. If $\tilde{v}_{\min}^{(1)} \geq 8D^*$, then a *tEFX* allocation exists.*

Proof. Let A be the allocation of Algorithm 1, and let $i, i' \in \mathcal{N}$ be any pair of agents such that $\tilde{v}_1^{(1)}(A_i) \geq \tilde{v}_1^{(1)}(A_{i'})$. By the definition of TV, $|\tilde{v}_1^{(1)}(A_i) - \tilde{v}_i^{(1)}(A_i)| \leq 2D_{TV}(\tilde{v}_1^{(1)} \parallel \tilde{v}_i^{(1)}) \leq 2D^*$, and $|v_1(A_{i'}) - v_i(A_{i'})| \leq 2D^*$. Combining the inequalities, we have that $\tilde{v}_i^{(1)}(A_{i'}) - \tilde{v}_i^{(1)}(A_i) = \tilde{v}_i^{(1)}(A_{i'}) + (\tilde{v}_1^{(1)}(A_{i'}) - \tilde{v}_1^{(1)}(A_i)) + (-\tilde{v}_1^{(1)}(A_i) + \tilde{v}_1^{(1)}(A_i)) - \tilde{v}_i^{(1)}(A_i) \leq 4D^* \leq \frac{\tilde{v}_{\min}^{(1)}}{2}$; therefore, since envy is scale-free, the removal of any item from $A_{i'}$ eliminates i 's envy.

Next, let j be the last item i received in Algorithm 1; then, it must be that $\tilde{v}_1^{(1)}(A_i) - \tilde{v}_{1,j}^{(1)} \leq \tilde{v}_1^{(1)}(A_{i'})$. By the definition of TV, we have $|\tilde{v}_1^{(1)}(A_i) - \tilde{v}_{i'}^{(1)}(A_i)| \leq 2D^*$ and $|\tilde{v}_1^{(1)}(A_{i'}) - \tilde{v}_{i'}^{(1)}(A_{i'})| \leq 2D^*$. Combining, we have that $\tilde{v}_{i'}^{(1)}(A_i) - \tilde{v}_{i'}^{(1)}(A_{i'}) \leq 4D^* + \tilde{v}_{1,j}^{(1)}$. Let $j^* \in \operatorname{argmin}_{\ell \in A_i} \tilde{v}_{i',\ell}^{(1)}$. By the definition of TV, we have $|\tilde{v}_{i',j^*}^{(1)} - \tilde{v}_{1,j^*}^{(1)}| \leq 2D^*$ and $|\tilde{v}_{i',j}^{(1)} - \tilde{v}_{1,j}^{(1)}| \leq 2D^*$. Since $\tilde{v}_{i',j^*}^{(1)} \leq \tilde{v}_{i',j}^{(1)}$ ($j \in A_i$ and $j^* \in \operatorname{argmin}_{\ell \in A_i} \tilde{v}_{i',\ell}^{(1)}$), and $\tilde{v}_{1,j}^{(1)} \leq \tilde{v}_{1,j^*}^{(1)}$ (j^* was included before j in the bundle A_i), we have that $\tilde{v}_{1,j}^{(1)} \leq 4D^* + \tilde{v}_{i',j^*}^{(1)}$. Combining, we have that $\tilde{v}_{i'}^{(1)}(A_i) - \tilde{v}_{i'}^{(1)}(A_{i'}) \leq 8D^* + \tilde{v}_{i',j^*}^{(1)} \leq \tilde{v}_{min}^{(1)} + \tilde{v}_{i',j^*}^{(1)}$. Therefore, transferring any item from A_i to $A_{i'}$ eliminates the envy of i' . $\square$