Fair Division Through Information Withholding

Hadi Hosseini[1], Sujoy Sikdar[2], Rohit Vaish[3], Hejun Wang[3], Lirong Xia[3]
[1]Rochester Institute of Technology, [2]Washington University St. Louis, [3]Rensselaer Polytechnic Institute
hho@cs.rit.edu, sujoy@wustl.edu, vaishr2@rpi.edu, wangj38@rpi.edu, xial@cs.rpi.edu

Abstract

Envy-freeness up to one good (EF1) is a well-studied fairness notion for indivisible goods that addresses pairwise envy by the removal of at most one good. In the worst case, each pair of agents might require the (hypothetical) removal of a different good, resulting in a weak *aggregate* guarantee. We study allocations that are nearly envy-free in aggregate, and define a novel fairness notion based on *information withholding*. Under this notion, an agent can withhold (or hide) some of the goods in its bundle and reveal the remaining goods to the other agents. We observe that in practice, envy-freeness can be achieved by withholding only a small number of goods overall. We show that finding allocations that withhold an optimal number of goods is computationally hard even for highly restricted classes of valuations. In contrast to the worst-case results, our experiments on synthetic and real-world preference data show that existing algorithms for finding EF1 allocations withhold a close-to-optimal amount of information.

1 Introduction

When dividing discrete objects, one often strives for a fairness notion called *envy-freeness* (Foley 1967), under which no agent prefers the allocation of another agent to its own. Such outcomes might not exist in general (even with only two agents and a single indivisible good), motivating the need for approximations. Among the many approximations of envy-freeness proposed in the literature (Lipton et al. 2004; Budish 2011; Nguyen and Rothe 2014; Caragiannis et al. 2016), the one that has found impressive practical appeal is *envy-freeness up to one good* (EF1). In an EF1 allocation, agent a can envy agent b as long as there is some good in b's bundle whose removal makes the envy go away. It is known that an EF1 allocation always exists and can be computed in polynomial time (Markakis 2017).

On closer scrutiny, however, we find that EF1 is not as strong as one might think: In the worst case, an EF1 allocation could entail the (hypothetical) removal of *every* good, because the elimination of each agent's envy may require the removal of a different good. To see this, consider an instance with six goods $g_1, \ldots, g_6$ and three agents a_1, a_2, a_3 whose (additive) valuations are as follows:

	g_1	g_2	g_3	g_4	g_5	g_6
a_1	$\textcircled{1}$	$\textcircled{1}$	$\underline{4}$	1	1	$\underline{4}$
a_2	1	$\underline{4}$	$\textcircled{1}$	$\textcircled{1}$	$\underline{4}$	1
a_3	$\underline{4}$	1	1	$\underline{4}$	$\textcircled{1}$	$\textcircled{1}$

Observe that the allocation shown via circled goods is EF1, since any pairwise envy can be addressed by removing an underlined good. However, each pair of agents requires the removal of a *different* good (e.g., a_1's envy towards a_2 is addressed by removing g_3 whereas a_3's envy towards a_2 is addressed by removing g_4, and so on), resulting in a weak approximation overall (since all goods need to be removed over all pairs of agents).

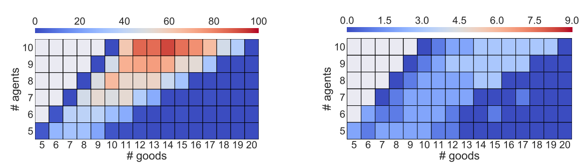
The above example shows that EF1, on its own, is too *coarse* to distinguish between allocations that remove a *large* number of goods (such as the one with circled entries) and those that remove only a *few* (such as the one with underlined entries, which, in fact, is envy-free). This limitation highlights the need for a fairness notion that (a) can distinguish between allocations in terms of their *aggregate* approximation, and (b) retains the "up to one good" style approximation of EF1 that has proven to be practically useful (Goldman and Procaccia 2014). Our work aims to fill this important gap.

We propose a new fairness notion called *envy-freeness up to k hidden goods* (HEF-k), defined as follows: Say there are n agents, m goods, and an allocation $A = (A_1, \ldots, A_n)$. Suppose there is a set S of k goods (called the *hidden* set) such that each agent i withholds the goods in $A_i \cap S$ (i.e., the hidden goods owned by i) and only discloses the goods in $A_i \setminus S$ to the other agents. Any other agent $h \neq i$ only observes the goods disclosed by i (i.e., those in $A_i \setminus S$), and its valuation for i's bundle is therefore $v_h(A_i \setminus S)$ instead of $v_h(A_i)$. Additionally, agent h's valuation for its own bundle is $v_h(A_h)$ (and not $v_h(A_h \setminus S)$) because it can observe its own hidden goods. If, under the disclosed allocation, no agent prefers the bundle of any other agent (i.e., if $v_h(A_h) \geq v_h(A_i \setminus S)$ for every pair of agents i, h), then we say that A is *envy-free up to k hidden goods* (HEF-k). In other words, by withholding the information about S, allocation A can be made free of envy.

Notice how HEF-k addresses the previous concerns: Like EF1, HEF-k is a relaxation of envy-freeness that is defined in terms of the *number of goods*. However, unlike EF1, HEF-k offers a *precise quantification* of the extent of information that must be withheld in order to achieve envy-freeness.

Clearly, any allocation can be made envy-free by hiding all the goods (i.e., if $k = m$). The real strength of HEF-k lies in k being *small*; indeed, an HEF-0 allocation is envy-free. As we will demonstrate below, there are natural settings that admit HEF-k allocations with a small k (i.e., hide only a small number of goods) even when (exact) envy-freeness is unlikely.

Information Withholding is Meaningful in Practice. To understand the usefulness of HEF-k, we generated a synthetic dataset where we varied the number of agents n from 5 to 10, and the number of goods m from 5 to 20 (we ignore the cases where $m < n$). For every fixed n and m, we generated 100 instances with *binary* valuations. Specifically, for every agent i and every good j, the valuation $v_{i,j}$ is drawn i.i.d. from Bernoulli(0.7). Figure 1a shows the heatmap of the number of instances out of 100 that *do not* admit envy-free outcomes. Figure 1b shows the heatmap of the number of goods that must be hidden in the worst-case. That is, the color of each cell denotes the smallest k such that each of the corresponding 100 instances admits some HEF-k allocation.



(a) Heatmap of the fraction of instances that are not envy-free.

(b) Heatmap of the number of goods that must be hidden.

Figure 1: In both figures, each cell corresponds to 100 instances with binary valuations for a fixed number of goods m (on X-axis) and a fixed number of agents n (on Y-axis).

It is evident from Figure 1 that even in the regime where envy-free outcomes are unlikely (in particular, the red-colored cells in Figure 1a), there exist HEF-k allocations with $k \leq 3$ (the light blue-colored cells in Figure 1b). This observation, along with the foregoing discussion, shows that fairness through information withholding is a well-motivated approach towards approximate envy-freeness that yields promising existence results in practice.

Our Contributions We make contributions on three fronts.

- On the *conceptual* side, we propose a novel fairness notion called HEF-k as a fine-grained generalization of envy-freeness in terms of aggregate approximation.

- Our *theoretical* results (Section 4) show that computing HEF-k allocations is computationally hard even for highly restricted classes of valuations (Theorem 1 and Corollary 1). We show a similar result when HEF-k is coupled with Pareto optimality (Theorem 2). We also provide an alternative proof of NP-completeness of determining the existence of an envy-free allocation for binary valuations (Proposition 3).

- Our *experiments* show that HEF-k allocations with a small k often exist, even when envy-free allocations do not (Figure 1). We also compare several known algorithms for computing EF1 allocations on synthetic and real-world preference data, and find that the round-robin algorithm and a recent algorithm of Barman, Krishnamurthy, and Vaish (2018) withhold close-to-optimal amount of information, often hiding no more than three goods (Section 5).

2 Related Work

An emerging line of work in the fair division literature considers relaxations of envy-freeness by limiting the information available to the agents. Notably, Aziz et al. (2018) consider a setting where each agent is aware only of its own bundle and has no knowledge about the allocations of the other agents. They propose the notion of *epistemic envy-freeness* (EEF) under which each agent believes that an envy-free allocation of the remaining goods among the other agents is possible. Note that in EEF, each agent might consider a different hypothetical assignment of the remaining goods, and each of these could be significantly different from the *actual* underlying allocation. By contrast, under HEF-k, each agent evaluates its valuation with respect to the same (underlying) allocation. Chen and Shah (2017) study a related model where agents have probabilistic beliefs about the allocations of the other agents, and envy is defined in expectation. Chan et al. (2019) study a setting similar to Aziz et al. (2018) wherein each agent is unaware of the allocations of the other agents, with the guarantee that it does not get the worst bundle.

Another related line of work considers settings where the agents constitute a social network and can only observe the allocations of their neighbors (Abebe, Kleinberg, and Parkes 2017; Bei, Qiao, and Zhang 2017; Chevaleyre, Endriss, and Maudet 2017; Aziz et al. 2018; Beynier et al. 2018; Bredereck, Kaczmarczyk, and Niedermeier 2018). These works place an informational constraint on the *set of agents*, whereas our model restricts the *set of revealed goods* per agent.

Several other forms of fairness approximations have been proposed recently, such as by introducing side payments (Halpern and Shah 2019), permitting sharing of some goods (Sandomirskiy and Segal-Halevi 2019), or donating a small fraction of goods (Caragiannis, Gravin, and Huang 2019).

3 Preliminaries

Problem instance An *instance* $\mathcal{I} = \langle [n], [m], \mathcal{V} \rangle$ of the fair division problem is defined by a set of $n \in \mathbb{N}$ *agents* $[n] = \{1, 2, \ldots, n\}$, a set of $m \in \mathbb{N}$ *goods* $[m] = \{1, 2, \ldots, m\}$, and a *valuation profile* $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ that specifies the preferences of every agent $i \in [n]$ over each subset of the goods in $[m]$ via a *valuation function* $v_i : 2^{[m]} \to \mathbb{N} \cup \{0\}$. Notice that each agent's valuation for any subset of goods is assumed to be a non-negative integer. We will assume that the valuation functions are *additive*, i.e., for any $i \in [n]$ and $G \subseteq [m]$, $v_i(G) := \sum_{j \in G} v_i(\{j\})$, where $v_i(\emptyset) = 0$. We will write $v_{i,j}$ instead of $v_i(\{j\})$ for a singleton good $j \in [m]$. We say that an instance has *binary valuations* if for every $i \in [n]$ and every $j \in [m]$, $v_{i,j} \in \{0, 1\}$.

Allocation An *allocation* $A := (A_1, \ldots, A_n)$ refers to an n-partition of the set of goods $[m]$, where $A_i \subseteq [m]$ is the *bundle* allocated to agent i. Given an allocation A, the utility of agent $i \in [n]$ for the bundle A_i is $v_i(A_i) = \sum_{j \in A_i} v_{i,j}$.

Definition 1 (Envy-freeness). An allocation A is *envy-free* (EF) if for every pair of agents $i, h \in [n]$, $v_i(A_i) \geq v_i(A_h)$. An allocation A is *envy-free up to one good* (EF1) if for every pair of agents $i, h \in [n]$ such that $A_h \neq \emptyset$, there exists some good $j \in A_h$ such that $v_i(A_i) \geq v_i(A_h \setminus \{j\})$. An allocation A is *strongly envy-free up to one good* (sEF1) if for every agent $h \in [n]$ such that $A_h \neq \emptyset$, there exists a good $g_h \in A_h$ such that for all $i \in [n]$, $v_i(A_i) \geq v_i(A_h \setminus \{g_h\})$. The notions of EF, EF1, and sEF1 are due to Foley (1967), Budish (2011), and Conitzer et al. (2019), respectively.[1]

Definition 2 (Envy-freeness with hidden goods). An allocation A is said to be *envy-free up to k hidden goods* (HEF-k) if there exists a set $S \subseteq [m]$ of at most k goods such that for every pair of agents $i, h \in [n]$, we have $v_i(A_i) \geq v_i(A_h \setminus S)$. An allocation A is *envy-free up to k uniformly hidden goods* (uHEF-k) if there exists a set $S \subseteq [m]$ of at most k goods satisfying $|S \cap A_i| \leq 1$ for every $i \in [n]$ such that for every pair of agents $i, h \in [n]$, we have $v_i(A_i) \geq v_i(A_h \setminus S)$. We say that allocation A *hides* the goods in S and *reveals* the remaining goods. Notice that a uHEF-k allocation is also HEF-k but the converse is not necessarily true. Indeed, in Proposition 2, we will present an instance that, for some $k \in \mathbb{N}$, admits an HEF-k allocation but no uHEF-k allocation.

[1] A slightly weaker notion than EF1 was previously studied by Lipton et al. (2004). However, their algorithm can be shown to compute an EF1 allocation.

Remark 1. It follows from the definitions that an allocation is EF if and only if it is HEF-0. It is also easy to verify that an allocation is sEF1 if and only if it is uHEF-n. This is because the unique hidden good for every agent is also the one that is (hypothetically) removed under sEF1. Additionally, as discussed in Section 1, an EF1 allocation might not be uHEF-k for any $k \leq n$.

We say that allocation A is HEF *with respect to set* S if A becomes envy-free after hiding the goods in S, i.e., for every pair of agents $i, h \in [n]$, we have $v_i(A_i) \geq v_i(A_h \setminus S)$. We say that k goods *must be hidden* under A if A is HEF with respect to some set S such that $|S| = k$, and there is no set S' with $|S'| < k$ such that A is HEF with respect to S'.

Definition 3 (Pareto optimality). An allocation A is Pareto dominated by another allocation B if $v_i(B_i) \geq v_i(A_i)$ for every agent $i \in [n]$ with at least one of the inequalities being strict. A *Pareto optimal* (PO) allocation is one that is not Pareto dominated by any other allocation.

Definition 4 (EF1 algorithms). We will now describe four known algorithms for finding EF1 allocations that are relevant to our work.

Round-robin algorithm (`RoundRobin`): Fix a permutation σ of the agents. The `RoundRobin` algorithm cycles through the agents according to σ. In each round, an agent gets its favorite good from the pool of remaining goods.

Envy-graph algorithm (`EnvyGraph`): This algorithm, proposed by Lipton et al. (2004), works as follows: In each step, one of the remaining goods is assigned to an agent that is not envied by any other agent. The existence of such an agent is guaranteed by resolving cyclic envy relations (if any exists) in a combinatorial structure called the *envy-graph* of an allocation.

Fisher market-based algorithm (`Alg-EF1+PO`): This algorithm, due to Barman, Krishnamurthy, and Vaish (2018), uses local search and price-rise subroutines in a Fisher market associated with the fair division instance, and returns an EF1 and PO allocation. The bound on running time of this algorithm is pseudopolynomial (a polynomial in $v_{i,j}$ instead of $\log v_{i,j}$).

Maximum Nash Welfare solution (`MNW`): The *Nash social welfare* of an allocation A is defined as $\text{NSW}(A) := \left(\prod_{i \in [n]} v_i(A_i)\right)^{1/n}$. The `MNW` algorithm computes an allocation with the highest Nash social welfare (called a *Nash optimal* allocation). Caragiannis et al. (2016) showed that a Nash optimal allocation is both EF1 and PO.

Remark 2. Conitzer et al. (2019) observed that `RoundRobin`, `Alg-EF1+PO`, and `MNW` algorithms all satisfy sEF1. It is easy to see that `EnvyGraph` algorithm is also sEF1. However, note that among the above algorithms, only `MNW` and `Alg-EF1+PO` are known to also satisfy PO.[2] The allocations computed by all four algorithms have the property that there exists some agent that is not envied by any other agent. Indeed, `MNW` and `Alg-EF1+PO` are both PO and therefore cannot have cyclic envy relations, and `RoundRobin` and `EnvyGraph` algorithms have this property by design. For such an agent (not necessarily the same agent for all four algorithms), no good needs to be removed under sEF1. Therefore, from Remark 1, all these algorithms are also envy-free up to $n-1$ uniformly hidden goods, or uHEF-$(n-1)$.

Proposition 1. *Given an instance with additive valuations, a* uHEF-$(n-1)$ *allocation always exists and can be computed in polynomial time, and a* uHEF-$(n-1)+$ PO *allocation always exists and can be computed in pseudopolynomial time.*

Remark 3. Note that for any $k < n-1$, an HEF-k allocation might fail to exist. Indeed, with n agents that have identical and positive valuations for $m = n-1$ goods, some agent will surely miss out and force the allocation to hide all $n-1$ (i.e., $k+1$ or more) goods. Therefore, the bound in Proposition 1 for uHEF-k (and hence, for HEF-k) is tight in terms of k.

Relevant Computational Problems

Definition 5 formalizes the decision problem of whether a given instance admits an HEF-k allocation.

Definition 5 (HEF-k-Existence). Given an instance $\mathcal{I}$, does there exist an allocation A and a set $S \subseteq [m]$ of at most k goods such that A is HEF w.r.t. S?

Notice that a certificate for HEF-k-Existence consists of an allocation A as well as a set S of at most k hidden goods. Another relevant computational question involves checking whether a given allocation A is HEF with respect to some set $S \subseteq [m]$ of at most k goods.

Definition 6 (HEF-k-Verification). Given an instance $\mathcal{I}$ and an allocation A, does there exist a set $S \subseteq [m]$ of k goods such that A is HEF w.r.t. S?

For additive valuations, both HEF-k-Existence and HEF-k-Verification are in NP. The next problem pertains to the existence of envy-free allocations.

Definition 7 (EF-Existence). Given an instance $\mathcal{I}$, does there exist an envy-free allocation for $\mathcal{I}$?

EF-Existence is known to be NP-complete (Lipton et al. 2004). From Remark 1, it follows that HEF-k-Existence is NP-complete when $k = 0$ for additive valuations.

4 Theoretical Results

We will now present our theoretical results concerning the existence and computation of HEF-k and uHEF-k allocations. We first show that uHEF-k is a strictly more demanding notion than HEF-k (Proposition 2).

[2]It is also known that `RoundRobin` and `EnvyGraph` fail to satisfy PO; see, e.g., (Conitzer, Freeman, and Shah 2017).

Proposition 2. *There exists an instance $\mathcal{I}$ that, for some fixed $k \in \mathbb{N}$, admits an* HEF-k *allocation but no* uHEF-k *allocation.*

Proof. Consider the fair division instance $\mathcal{I}$ with five agents $a_1, \ldots, a_5$ and six goods $g_1, \ldots, g_6$ shown in Table 1. Observe that the allocation $A = (A_1, \ldots, A_5)$ with $A_1 = \{g_1, g_2\}$, $A_2 = \{g_3\}$, $A_3 = \{g_4\}$, $A_4 = \{g_5\}$, $A_5 = \{g_6\}$ satisfies HEF-2 with respect to the set $S = \{g_1, g_2\}$.

	g_1	g_2	g_3	g_4	g_5	g_6
a_1	1	1	2	0	0	0
a_2	1	1	2	0	0	0
a_3	10	10	1	1	1	1
a_4	10	10	1	1	1	1
a_5	10	10	1	1	1	1

Table 1: The instance used in the proof of Proposition 2.

We will show that $\mathcal{I}$ does not admit a uHEF-2 allocation. Suppose, for contradiction, that there exists an allocation B satisfying uHEF-2. Then, B must hide g_1 and g_2 (otherwise, at least one of a_3, a_4 or a_5 will envy the owner(s) of these goods). Thus, in particular, the good g_3 must be revealed by B. Assume, without loss of generality, that g_3 is *not* assigned to a_1 in B (otherwise, a similar argument can be carried out for a_2). Then, B must assign both g_1 and g_2 to a_1 (so that a_1 does not envy the owner of g_3). However, this violates the one-hidden-good-per-agent property of uHEF-k, which is a contradiction. ☐

Recall from Section 3 that HEF-k-Existence is NP-complete when $k = 0$. This still leaves open the question whether HEF-k-Existence is NP-complete for *any* fixed $k \in \mathbb{N}$. Our next result (Theorem 1) shows that this is indeed the case, even under the restricted setting of *identical* valuations (i.e., for every $j \in [m]$, $v_{i,j} = v_{h,j}$ for every $i, h \in [n]$).

Theorem 1 (Hardness of HEF-k-Existence). *For any fixed $k \in \mathbb{N}$,* HEF-k-Existence *is* NP-complete *even for identical valuations.*

Proof. We will show a reduction from Partition, which is known to be NP-complete (Garey and Johnson 1979). An instance of Partition consists of a multiset $X = \{x_1, x_2, \ldots, x_n\}$ with $x_i \in \mathbb{N}$ for all $i \in [n]$. The goal is to determine whether there exists $Y \subset X$ such that $\sum_{x_i \in Y} x_i = \sum_{x_i \in X \setminus Y} x_i = T$, where $T \coloneqq \frac{1}{2} \sum_{x_i \in X} x_i$.

We will construct a fair division instance with $k+3$ agents $a_1, \ldots, a_{k+3}$ and $n+k+1$ goods. The goods are classified into $n+1$ *main goods* $g_1, \ldots, g_{n+1}$ and k *dummy goods* $d_1, \ldots, d_k$. The (identical) valuations are defined as follows: Every agent values the goods $g_1, \ldots, g_n$ at $x_1, \ldots, x_n$ respectively; the good g_{n+1} at T, and each dummy good at $4T$.

($\Rightarrow$) Suppose Y is a solution of Partition. Then, an HEF-k allocation can be constructed as follows: Assign the main goods corresponding to the set Y to agent a_1 and those corresponding to $X \setminus Y$ to agent a_2. The good g_{n+1} is assigned to agent a_3. Each of the remaining k agents is assigned a unique dummy good. Note that every agent in the set $\{a_1, a_2, a_3\}$ envies every agent in the set $\{a_4, \ldots, a_{k+3}\}$, and these are the only pairs of agents with non-zero envy. Therefore, the allocation can be made envy-free by hiding the k dummy goods, i.e., the allocation is HEF with respect to the set $\{d_1, \ldots, d_k\}$.

($\Leftarrow$) Now suppose there exists an HEF-k allocation A. Since there are k dummy goods and $k+3$ agents, there must exist at least three agents that do not receive any dummy good in A. Without loss of generality, let these agents be a_1, a_2 and a_3 (otherwise, we can reindex). We claim that all dummy goods must be hidden under A. Indeed, agent a_1 does not receive any dummy good, and therefore its maximum possible valuation can be $v(g_1 \cup \cdots \cup g_{n+1}) = 3T < v(d_j)$ for any dummy good d_j. If some dummy good d_j is not hidden, then a_1 will envy the owner of d_j, contradicting HEF-k. Therefore, all dummy goods must be hidden, and since there are k such goods, these are the only ones that can be hidden.

The above observation implies that the good g_{n+1} must be revealed by A. Furthermore, g_{n+1} must be assigned to one of a_1, a_2 or a_3 (otherwise, by pigeonhole principle, one of these agents will have valuation at most $\frac{2T}{3}$ and will envy the owner of g_{n+1}). If g_{n+1} is assigned to a_3, then the remaining main goods $g_1, \ldots, g_n$ must be divided between a_1 and a_2 such that $v(A_1) \geq T$ and $v(A_2) \geq T$. This gives a partition of the set X. □

Another commonly used preference restriction is that of *binary* valuations (i.e., for every $i \in [n]$ and $j \in [m], v_{i,j} \in \{0, 1\}$). We note that even under this restriction, HEF-k-Existence remains NP-complete when $k = 0$ (Corollary 1). This observation follows from a result of Aziz et al. (2015), who showed that determining the existence of an envy-free allocation is NP-complete even for binary valuations (Proposition 3). We provide an alternative proof of this statement in the full version of the paper (Hosseini et al. 2019).

Proposition 3 (Aziz et al. 2015; Theorem 11)**.** EF-Existence *is* NP-complete *even for binary valuations.*

Corollary 1. *For* $k = 0$*,* HEF-k-Existence *is* NP-complete *even for binary valuations.*

Proposition 3 is also useful in establishing the computational hardness of finding an HEF-k+PO allocation. Note that unlike Corollary 1, Theorem 2 holds for any fixed $k \in \mathbb{N}$.

Theorem 2 (Hardness of HEF-k+PO). *Given any instance $\mathcal{I}$ with binary valuations and any fixed $k \in \mathbb{N} \cup \{0\}$, it is* NP-hard *to determine if $\mathcal{I}$ admits an allocation that is envy-free up to k hidden goods* (HEF-k) *and Pareto optimal* (PO).

Proof. (Sketch) Starting from any instance of EF-Existence with binary valuations (Proposition 3), we add to it k new goods and $k+1$ new agents such that all new goods are approved by all new agents (and no one else). Also, the new agents have zero value for the existing goods. In the forward direction, an arbitrary allocation of new goods among the new agents works. In the reverse direction, PO forces each new (respectively, existing) good to be assigned among new (respectively, existing) agents only. The imbalance between new agents and new goods means that all (and only) the new goods must be hidden. Then, the restriction of the HEF-k allocation to the existing agents/goods gives the desired EF allocation. □

We will now proceed to analyzing the computational complexity of HEF-k-Verification. Here, we show a hardness-of-approximation result (Theorem 3). Note that HEF-k-Verification is stated as a decision problem. However, one can consider an approximation version of this problem as follows: A c-approximation algorithm for HEF-k-Verification is one that, given any fair division instance, computes a set of goods of size at most $c \cdot k^{\text{opt}}$, where k^{opt} is the size of the smallest hidden set for the given instance. Under this definition, Theorem 3 can be interpreted as follows: Given any $\varepsilon > 0$, there is no polynomial-time $(1-\varepsilon). \ln E$-approximation algorithm for HEF-k-Verification, unless P=NP.

Theorem 3 (HEF-k-Verification inapproximability). *Given any $\varepsilon > 0$, it is NP-hard to approximate* HEF-k-Verification *to within $(1-\varepsilon) \cdot \ln E$ even for binary valuations, where E is the sum of all non-negative pairwise envy values in the given allocation.*

Proof. We will show a reduction from Hitting Set. An instance of Hitting Set consists of a finite set $X = \{x_1, \ldots, x_p\}$, a collection $\mathcal{F} = \{F_1, \ldots, F_q\}$ of subsets of X, and some $k \in \mathbb{N}$. The goal is to determine whether there exists $Y \subseteq X$, $|Y| \leq k$ that intersects every member of $\mathcal{F}$ (i.e., for every $F \in \mathcal{F}, Y \cap F \neq \emptyset$). It is known that given any $\varepsilon > 0$, it is NP-hard to approximate Hitting Set to within a factor $(1-\varepsilon) \cdot \ln |\mathcal{F}|$ (Dinur and Steurer 2014).

We will construct a fair division instance with $n = q + 1$ agents and $m = p + \sum_{i=1}^{q}(|F_i| - 1)$ goods. The agents are classified into q *dummy agents* $a_1, \ldots, a_q$ and one *main agent* a_{q+1}. The goods are classified into p *main goods* $g_1, \ldots, g_p$ and q distinct sets of dummy goods, where the i^{th} set consists of the goods $f^i_1, \ldots, f^i_{|F_i|-1}$.

The valuations are as follows: The main agent approves all the main goods, i.e., for all $j \in [p]$, $v_{q+1}(\{g_j\}) = 1$. Each dummy agent a_i approves the dummy goods in the i^{th} set as well as those main goods that intersect with F_i, i.e., for every $i \in [q], v_i(\{f^i_j\}) = 1$ for all $j \in [|F_i| - 1]$, and $v_i(\{g_j\}) = 1$ whenever $x_j \in F_i$. All other valuations are set to 0.

The input allocation $A = (A_1, \ldots, A_{q+1})$ is defined as follows: The main agent a_{q+1} is assigned all the main goods, i.e., $A_{q+1} \coloneqq \{g_1, \ldots, g_p\}$. For every $i \in [q]$, the dummy agent a_i is assigned the $|F_i| - 1$ dummy goods in the i^{th} set, i.e., $A_i \coloneqq \{f_1^i, \ldots, f_{|F_i|-1}^i\}$. Note that in the allocation A, each dummy agent envies the main agent by one approved good, and these are the only pairs of agents with envy. Finally, given any allocation A, we define the *aggregate envy* in A as the sum of all non-negative pairwise envy values, i.e.,

$$E \coloneqq \textstyle\sum_{h \in [n]} \sum_{i \neq h} \max\{0, v_i(A_h) - v_i(A_i)\}.$$

($\Rightarrow$) Suppose $Y \subseteq X$, $|Y| \leq k$ is solution of the Hitting Set instance. We claim that the allocation A is HEF with respect to the set $S \coloneqq \{g_j : x_j \in Y\}$ with $|S| \leq k$. Indeed, since S is induced by a hitting set, each dummy agent approves at least one good in S. Therefore, by hiding the goods in S, the envy from the dummy agents can be eliminated.

($\Leftarrow$) Now suppose there exists $S \subseteq [m]$, $|S| \leq k$ such that A is HEF with respect to S. Then, for every $i \in [q]$, the set S must contain at least one good that is approved by the dummy agent a_i (otherwise A will not be envy-free after hiding the goods in S). It is easy to see that the set $Y \coloneqq \{x_j : g_j \in S\}$ constitutes the desired hitting set of cardinality at most k.

Finally, to show the hardness-of-approximation, notice that the aggregate envy in A is q because each dummy agent envies the main agent by one unit of utility. The claim now follows by substituting $|\mathcal{F}| = q = E$ in the inapproximability result of Hitting Set stated above. □

Our next result (Theorem 4) provides an approximation algorithm that (nearly) matches the hardness-of-approximation result in Theorem 3. We remark that the algorithm in Theorem 4 applies to *any* instance with additive and possibly non-binary valuations.

Theorem 4 (**Approximation algorithm**). *There is a polynomial-time algorithm that, given as input any instance of* HEF-k-Verification*, finds a set* $S \subseteq [m]$ *with* $|S| \leq k^{\text{opt}} \cdot \ln E + 1$ *such that the given allocation is* HEF *with respect to* S*. Here,* E *and* k^{opt} *denote the aggregate envy and the number of goods that must be hidden under the given allocation, respectively.*

The proof of Theorem 4 is available in the full version (Hosseini et al. 2019), but a brief idea is as follows: For any set $S \subseteq [m]$, define the *residual envy function* $f : 2^{[m]} \to \mathbb{R}$ so that $f(S)$ is the aggregate envy in allocation A after hiding the goods in S. That is,

$$f(S) \coloneqq \textstyle\sum_{h \in [n]} \sum_{i \neq h} \max\{0, v_i(A_h \setminus S) - v_i(A_i)\}.$$

The relevant observation is that f is *supermodular*. Given this observation, the approximation guarantee in Theorem 4 can be obtained by the standard greedy algorithm for submodular maximization, or, equivalently, supermodular minimization (Nemhauser, Wolsey, and Fisher 1978); see Algorithm 1.

ALGORITHM 1: Greedy Approximation Algorithm for HEF-k-Verification

Input: An instance $\langle [n], [m], \mathcal{V} \rangle$ and an allocation A.
Output: A set $S \subseteq [m]$.

Initialize $S = \emptyset$.
while $f(S) \geq 1$ **do**
  Set $j' \leftarrow \arg\max_{j \in [m] \setminus S} f(S) - f(S \cup \{j\})$
    ▷ `tiebreak lexicographically`
  Update $S \leftarrow S \cup \{j'\}$
return S

5 Experimental Results

We have seen that the worst-case computational results for HEF-k, even in highly restricted settings, are largely negative (Section 4). In this section, we will examine whether the known algorithms for computing approximately envy-free allocations—in particular, the four EF1 algorithms described in Definition 4 in Section 3—can provide meaningful approximations to HEF-k in practice. Recall from Remark 2 that all four discussed algorithms—`RoundRobin`, `MNW`, `Alg-EF1+PO`, and `EnvyGraph`—satisfy uHEF-$(n-1)$.

We evaluate each algorithm in terms of (a) its *regret* (defined below), and (b) the *number of goods that the algorithm must hide*. Given an instance $\mathcal{I}$ and an allocation A, let $\kappa(A, \mathcal{I})$ denote the number of goods that must be hidden under A. The *regret* of allocation A is the number of extra goods that must be hidden under A compared to the optimal. That is, $\texttt{reg}(A, \mathcal{I}) \coloneqq \kappa(A, \mathcal{I}) - \min_B \kappa(B, \mathcal{I})$. Similarly, given an algorithm Alg, the regret of Alg is given by $\texttt{reg}(\textsc{Alg}(\mathcal{I}), \mathcal{I})$, where $\textsc{Alg}(\mathcal{I})$ is the allocation returned by Alg for the input instance $\mathcal{I}$. Note that the regret can be large due to the suboptimality of an algorithm, but also due to the size of the instance. To negate the effect of the latter, we normalize the regret value by $n - 1$, which is the worst-case upper bound on the number of hidden goods for all four algorithms of interest.

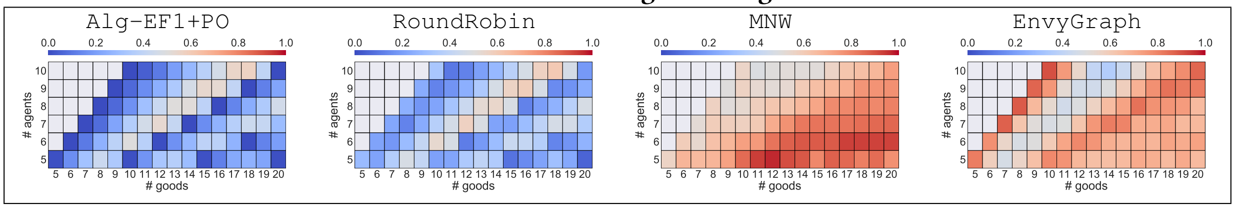
Experiments on Synthetic Data

The setup for synthetic experiments is similar to that used in Figure 1. Specifically, the number of agents, n, is varied from 5 to 10, and the number of goods, m, is varied from 5 to 20 (we ignore the cases where $m < n$). For every fixed n and m, we generated 100 instances with *binary* valuations drawn i.i.d. from Bernoulli distribution with parameter 0.7 (i.e., $v_{i,j} \sim \text{Ber}(0.7)$). Table 2 shows the heatmaps of the normalized regret (averaged over 100 instances) and the number of goods that must be hidden (averaged over non-EF instances, i.e., whenever $k \geq 1$) for all four algorithms.[3]

It is clear that `Alg-EF1+PO` and `RoundRobin` algorithms have a superior performance than `MNW` and

[3] The full version (Hosseini et al. 2019) provides additional results for $v_{i,j} \sim \text{Ber}(0.7)$, and $v_{i,j} \sim \text{Ber}(0.5)$.

Normalized average-case regret



Number of goods that must be hidden on average (averaged over non-EF instances only)

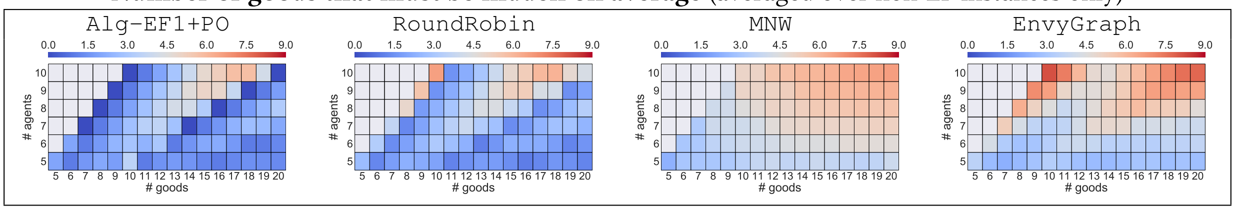


Table 2: Results for synthetic data.

EnvyGraph. In particular, both Alg-EF1+PO and RoundRobin have small normalized regret, suggesting that they hide close-to-optimal number of goods. Additionally, the number of hidden goods itself is small for these algorithms (in most cases, no more than *three* goods need to be hidden), suggesting that the worst-case bound of $n-1$ is unlikely to arise in practice. Overall, our experiments suggest that Alg-EF1+PO and RoundRobin can achieve useful approximations to HEF-k in practice, especially in comparison to MNW and EnvyGraph.[4]

Experiments on Real-World Data

For experiments with real-world data, we use the data from the popular fair division website *Spliddit* (Goldman and Procaccia 2014). The Spliddit data has 2212 instances in total, where the number of agents n varies between 3 and 10, and the number of goods $m \geq n$ varies between 3 and 93. Unlike the synthetic data, the distribution of instances here is rather uneven (see the full version online); in fact, 1821 of the 2212 instances have $n = 3$ agents and $m = 6$ goods. Therefore, instead of using heatmaps, we compare the algorithms in terms of their normalized regret (averaged over the entire dataset) and the cumulative distribution function of the hidden goods (see Figure 2).

Figure 2 presents an interesting twist: MNW is now the best performing algorithm, closely followed by RoundRobin and Alg-EF1+PO. For any fixed k, the fraction of instances for which these three algorithms compute an HEF-k allocation is also nearly identical. As can be observed, these algorithms almost never need to hide more than *three* goods. By contrast, EnvyGraph has the largest regret and significantly worse cumulative performance. Therefore, once again,

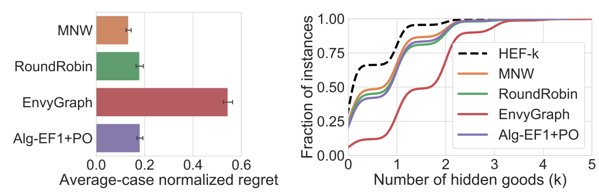


Figure 2: Results for Spliddit data.

Alg-EF1+PO and RoundRobin algorithms perform competitively with the optimal solution, making them attractive options for achieving fair outcomes without withholding too much information.

6 Future Work

The asymptotic existence of envy-free allocations has been studied by Dickerson et al. (2014) and Manurangsi and Suksompong (2019). Analyzing the asymptotic behavior of HEF-k allocations is an interesting direction for future work. Exploring the connection with other recently proposed relaxations that involve discarding goods (Caragiannis, Gravin, and Huang 2019) or sharing a small subset of goods (Sandomirskiy and Segal-Halevi 2019) might also be interesting.

Acknowledgments

We thank the anonymous reviewers for their helpful comments. We are grateful to Ariel Procaccia and Nisarg Shah for sharing with us the data from Spliddit, and to Haris Aziz for bringing to our attention the proof of EF-Existence for binary valuations in (Aziz et al. 2015). RV thanks Rupert Freeman, Nick Gravin, and Neeldhara Misra for very helpful discussions and several useful suggestions for improving the presentation of the paper. LX acknowledges NSF #1453542 and #1716333, and HH acknowledges NSF #1850076 for support.

[4]In the full version of the paper (Hosseini et al. 2019), we provide two families of instances where the normalized worst-case regret of MNW is large.

References

Abebe, R.; Kleinberg, J.; and Parkes, D. C. 2017. Fair Division via Social Comparison. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*, 281–289.

Aziz, H.; Gaspers, S.; Mackenzie, S.; and Walsh, T. 2015. Fair Assignment of Indivisible Objects under Ordinal Preferences. *Artificial Intelligence* 227:71–92.

Aziz, H.; Bouveret, S.; Caragiannis, I.; Giagkousi, I.; and Lang, J. 2018. Knowledge, Fairness, and Social Constraints. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 4638–4645.

Barman, S.; Krishnamurthy, S. K.; and Vaish, R. 2018. Finding Fair and Efficient Allocations. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, 557–574.

Bei, X.; Qiao, Y.; and Zhang, S. 2017. Networked Fairness in Cake Cutting. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 3632–3638.

Beynier, A.; Chevaleyre, Y.; Gourvès, L.; Lesca, J.; Maudet, N.; and Wilczynski, A. 2018. Local Envy-Freeness in House Allocation Problems. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, 292–300.

Bredereck, R.; Kaczmarczyk, A.; and Niedermeier, R. 2018. Envy-Free Allocations Respecting Social Networks. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, 283–291.

Budish, E. 2011. The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. *Journal of Political Economy* 119(6):1061–1103.

Caragiannis, I.; Kurokawa, D.; Moulin, H.; Procaccia, A. D.; Shah, N.; and Wang, J. 2016. The Unreasonable Fairness of Maximum Nash Welfare. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, 305–322.

Caragiannis, I.; Gravin, N.; and Huang, X. 2019. Envy-Freeness Up to Any Item with High Nash Welfare: The Virtue of Donating Items. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, 527–545. ACM.

Chan, H.; Chen, J.; Li, B.; and Wu, X. 2019. Maximin-Aware Allocations of Indivisible Goods. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 137–143.

Chen, Y., and Shah, N. 2017. Ignorance is Often Bliss: Envy with Incomplete Information. Technical report.

Chevaleyre, Y.; Endriss, U.; and Maudet, N. 2017. Distributed Fair Allocation of Indivisible Goods. *Artificial Intelligence* 242:1–22.

Conitzer, V.; Freeman, R.; Shah, N.; and Vaughan, J. W. 2019. Group Fairness for Indivisible Good Allocation. In *Thirty-Third AAAI Conference on Artificial Intelligence*, 1853–1860.

Conitzer, V.; Freeman, R.; and Shah, N. 2017. Fair Public Decision Making. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, 629–646. ACM.

Dickerson, J. P.; Goldman, J.; Karp, J.; Procaccia, A. D.; and Sandholm, T. 2014. The Computational Rise and Fall of Fairness. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 1405–1411.

Dinur, I., and Steurer, D. 2014. Analytical Approach to Parallel Repetition. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, 624–633.

Foley, D. 1967. Resource Allocation and the Public Sector. *Yale Economic Essays* 45–98.

Garey, M. R., and Johnson, D. S. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co.

Goldman, J., and Procaccia, A. D. 2014. Spliddit: Unleashing Fair Division Algorithms. *ACM SIGecom Exchanges* 13(2):41–46.

Halpern, D., and Shah, N. 2019. Fair Division with Subsidy. In *International Symposium on Algorithmic Game Theory*, 374–389. Springer.

Hosseini, H.; Sikdar, S.; Vaish, R.; Wang, J.; and Xia, L. 2019. Fair Division through Information Withholding. *arXiv preprint arXiv:1907.02583.*

Lipton, R. J.; Markakis, E.; Mossel, E.; and Saberi, A. 2004. On Approximately Fair Allocations of Indivisible Goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, 125–131.

Manurangsi, P., and Suksompong, W. 2019. When Do Envy-Free Allocations Exist? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 2109–2116.

Markakis, E. 2017. Approximation algorithms and hardness results for fair division with indivisible goods. In Endriss, U., ed., *Trends in Computational Social Choice.* AI Access. chapter 12, 231–247.

Nemhauser, G. L.; Wolsey, L. A.; and Fisher, M. L. 1978. An Analysis of Approximations for Maximizing Submodular Set Functions–I. *Mathematical Programming* 14(1):265–294.

Nguyen, T. T., and Rothe, J. 2014. Minimizing Envy and Maximizing Average Nash Social Welfare in the Allocation of Indivisible Goods. *Discrete Applied Mathematics* 179:54–68.

Sandomirskiy, F., and Segal-Halevi, E. 2019. Fair Division with Minimal Sharing. *arXiv preprint arXiv:1908.01669.*