Fair Division through Information Withholding*

Paper #:18

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 general, 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 our 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. On our way, we show that for binary valuations, finding an envy-free allocation is NP-complete—somewhat surprisingly, this fundamental question was unresolved prior to our work. 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 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 [Lipton et al., 2004].

A closer scrutiny, however, reveals 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. To see this, consider an instance with six goods $g_1, \dots, g_6$ 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	1	1	$\underline{4}$	1	1	$\underline{4}$
a_2	1	$\underline{4}$	1	1	$\underline{4}$	1
a_3	$\underline{4}$	1	1	$\underline{4}$	1	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 in aggregate (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 drawback 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 so useful in practice [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 $A = (A_1, \dots, A_n)$. 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.

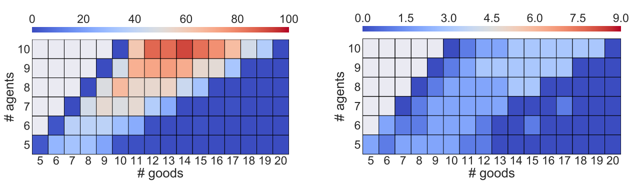
*Full version available at: https://gofile.io/?c=9dBhkZ

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.

Of course, any allocation can be made envy-free by hiding all the goods (i.e., if $k = m$). The true 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, $v_{i,j}$ 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 areas in Figure 1a), there exist HEF-k allocations with $k \leq 3$ (the light blue cells in Figure 1b). The above experiment, 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 3) 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 show that finding an envy-free allocation is NP-complete even for binary valuations (Lemma 1). Surprisingly, this fundamental problem was open prior to our work.

• 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 et al. [2018] withhold close-to-optimal information, often hiding no more than three goods (Section 4).

A discussion of related work is available in the full version.

2 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* $v_i : 2^{[m]} \to \mathbb{N} \cup \{0\}$. We will assume that the valuation functions are *additive*, i.e., for any $i \in [n]$ and $G \subseteq [m]$, $v_i(G) \coloneqq \sum_{j \in G} v_i(\{j\})$, 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* $A \coloneqq (A_1, \ldots, A_n)$ 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.

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 (see Proposition 2).

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.

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 (**EF**1 **algorithms**). We will now describe four known algorithms for finding EF1 allocations that are especially relevant to our work.

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

Envy-graph algorithm (`EnvyGraph`): This algorithm was proposed by Lipton et al. [2004] and works as follows: In each round, 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 in a combinatorial structure called the *envy-graph* of an allocation.

Fisher market-based algorithm (`Alg-EF1+PO`): This algorithm, due to Barman et al. [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 worst-case running time of this algorithm is pseudopolynomial (i.e., is 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. It is known that `MNW` and `Alg-EF1+PO` satisfy PO, while `RoundRobin` and `EnvyGraph` fail to satisfy PO (see, e.g., [Conitzer et al., 2017]). 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 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 $m = n-1$ goods that are identically valued by n agents, 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 (**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 with respect to S?

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 with respect to S?

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

3 Theoretical Results

We now briefly present our theoretical results. A more detailed discussion of the results is available in the full version. Our first result (Proposition 2) shows that uHEF-k is a strictly more demanding notion than HEF-k.

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

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

We show that even under the restriction of *binary* valuations, HEF-k-EXISTENCE remains NP-complete when $k = 0$ (Corollary 1). This follows from Lemma 1[1], which shows that for binary valuations, determining the existence of an envy-free allocation is NP-complete.

Lemma 1. EF-EXISTENCE *is* NP-complete *even for binary valuations.*

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

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).

We provide a hardness-of-approximation result (Theorem 3) for HEF-k-VERIFICATION. Here, the inapproximability factor is stated in terms of the aggregate envy, defined as follows: Given any allocation A, the *aggregate envy* in A is the sum of all pairwise envy values, i.e., $E_A := \sum_{h\in[n]} \sum_{i\neq h} \max\{0, v_i(A_h) - v_i(A_i)\}$.

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 aggregate envy in the given allocation.*

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.*

[1] We remark that our contribution is to show that EF-EXISTENCE remains NP-complete *even under binary valuations*; without this restriction, NP-completeness was already known [Lipton et al., 2004].

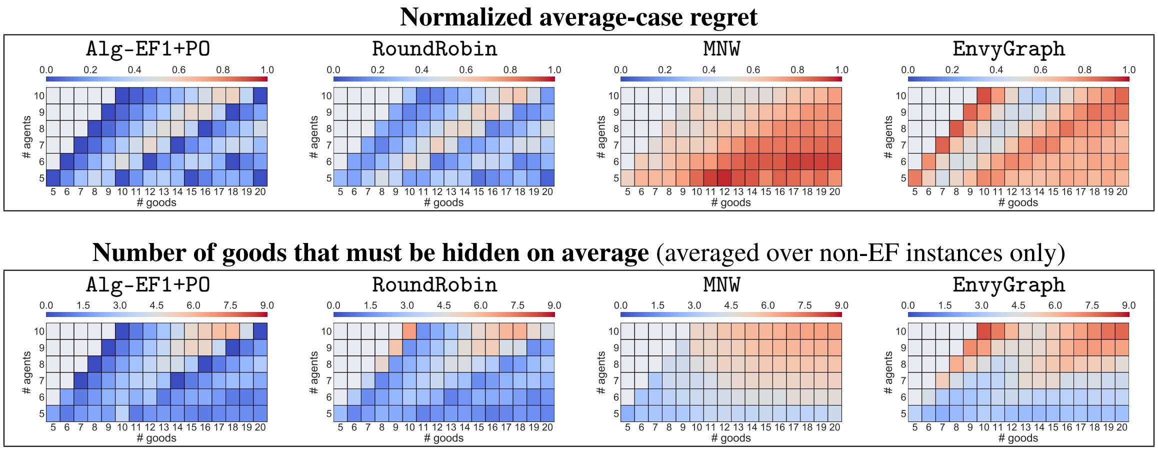


Table 1: *Results for synthetic data.*

4 Experimental Results

We have seen that the worst-case computational results for HEF-k, even in highly restricted settings, are mostly negative (Section 3). 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 2—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 $\mathcal{I}$, 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, $\texttt{reg}(A, \mathcal{I}) := \kappa(A, \mathcal{I}) - \min_B \kappa(B, \mathcal{I})$. Similarly, given an algorithm $\textsc{Alg}$, the regret of $\textsc{Alg}$ is given by $\texttt{reg}(\textsc{Alg}(\mathcal{I}), \mathcal{I})$, where $\textsc{Alg}(\mathcal{I})$ is the allocation returned by $\textsc{Alg}$ 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, as discussed above, is the worst-case upper bound on the number of hidden goods for all four algorithms of interest.

4.1 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 1 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.Additional experimental results and discussions are provided in the full version.

Our main observation is that both `Alg-EF1+PO` and `RoundRobin` have small normalized regret, suggesting that they hide close-to-optimal number of goods, suggest that `Alg-EF1+PO` and `RoundRobin` can achieve useful approximations to HEF-k in practice. 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.

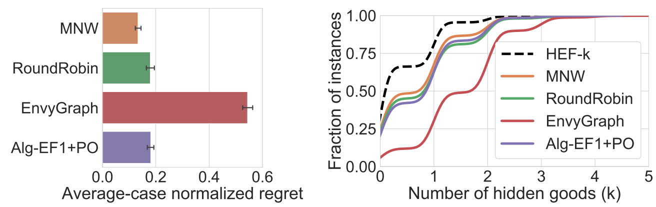


Figure 2: *Results for Spliddit data.*

4.2 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. Since the distribution of instances here is rather uneven (see Figure 3 in the supplementary material), 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. Therefore, once again, `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.

References

Siddharth Barman, Sanath Kumar Krishnamurthy, and Rohit Vaish. Finding Fair and Efficient Allocations. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 557–574, 2018.

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

Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D Procaccia, Nisarg Shah, and Junxing Wang. The Unreasonable Fairness of Maximum Nash Welfare. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*, pages 305–322, 2016.

Vincent Conitzer, Rupert Freeman, and Nisarg Shah. Fair Public Decision Making. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 629–646. ACM, 2017.

Vincent Conitzer, Rupert Freeman, Nisarg Shah, and Jennifer Wortman Vaughan. Group Fairness for Indivisible Good Allocation. In *Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19) (to appear)*, 2019.

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

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

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

Andreas Krause and Daniel Golovin. Submodular Function Maximization. In *Tractability: Practical Approaches to Hard Problems*, pages 71–104. Cambridge University Press, 2014.

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.

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

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