

# Envy-Freeness Up to Any Item with High Nash Welfare: The Virtue of Donating Items

IOANNIS CARAGIANNIS, University of Patras & CTI “Diophantus”

NICK GRAVIN, Shanghai University of Finance and Economics, ITCS

XIN HUANG, The Chinese University of Hong Kong

Several fairness concepts have been proposed recently in attempts to approximate envy-freeness in settings with indivisible goods. Among them, the concept of envy-freeness up to any item (EFX) is arguably the closest to envy-freeness. Unfortunately, EFX allocations are not known to exist except in a few special cases. We make significant progress in this direction. We show that for every instance with additive valuations, there is an EFX allocation of a subset of items with a Nash welfare that is at least half of the maximum possible Nash welfare for the original set of items. That is, after donating some items to a charity, one can distribute the remaining items in a fair way with high efficiency. This bound is proved to be best possible. Our proof is constructive and highlights the importance of maximum Nash welfare allocation. Starting with such an allocation, our algorithm decides which items to donate and redistributes the initial bundles to the agents, eventually obtaining an allocation with the claimed efficiency guarantee. The application of our algorithm to large markets, where the valuations of an agent for every item is relatively small, yields EFX with almost optimal Nash welfare. We also show that our algorithm can be modified to compute, in polynomial-time, EFX allocations that approximate optimal Nash welfare within a factor of at most  $2\rho$ , using a  $\rho$ -approximate allocation on input instead of the maximum Nash welfare one.

CCS Concepts: • **Theory of computation** → **Algorithmic game theory; Approximation algorithms analysis.**

Additional Key Words and Phrases: Fairness in allocation problems, indivisible items, envy-freeness, Nash welfare

## ACM Reference Format:

Ioannis Caragiannis, Nick Gravin, and Xin Huang. 2019. Envy-Freeness Up to Any Item with High Nash Welfare: The Virtue of Donating Items. In *ACM EC '19: ACM Conference on Economics and Computation (EC '19), June 24–28, 2019, Phoenix, AZ, USA*. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3328526.3329574>

## 1 INTRODUCTION

The agenda of fairly allocating items to agents has received much attention by the EconCS community recently. Most of the solution concepts with compelling properties such as envy-freeness and proportionality cannot always be attained when the items are indivisible, in contrast to the traditionally studied case of divisible items. Recent work put forth several promising fairness notions that adapt and “approximate” the definition of envy-freeness to the case of individual items and demonstrated that allocations that satisfy some of these notions are possible, with some minimum efficiency guarantees at the same time.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*EC '19, June 24–28, 2019, Phoenix, AZ, USA*

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6792-9/19/06...\$15.00

<https://doi.org/10.1145/3328526.3329574>

A very natural adaptation of envy-freeness for settings with indivisible items was defined by Budish [13] (but is also implicit in the earlier work of Lipton et al. [29]). An allocation of items is envy-free up to some item (EF1) if any possible envy of an agent for the allocation of another can be eliminated by removing *some* item from the envied bundle. It is well known that EF1 allocations can always be achieved, and recently Caragiannis et al. [15] showed that the Nash welfare maximizing allocation is Pareto-optimal and EF1. This is indeed a success story for the Nash welfare efficiency measure defined as the geometric mean of the agent valuations, a standard and well studied objective that is particularly well aligned with fair division applications.

For example, consider the problem of fairly dividing the inheritance of four items, a car, a ring, a painting, and a necklace, among three agents named Alice, Bob, and Carol. Agents have preferences for the items, which are expressed in the table at the right, where each number represents the amount in thousands of USD of how much each agent values the corresponding item. Clearly, there is no envy-free allocation

	c	r	p	n
Alice	10	9	4	6
Bob	10	6	9	4
Carol	10	4	6	9

in this example, because someone must get at least two items, and then one of the two other agents would have a higher total value for these items compared to what the agent received. The allocation which gives the ring to Alice, the car and the painting to Bob, and the necklace to Carol, is EF1 with the valuations of 9, 19, and 9, respectively. In this allocation, Bob does not envy anyone, while Alice and Carol would prefer to swap Bob's allocation with their own: Alice values the car and the painting in Bob's bundle for the total of 14, and Carol values them for 16. This is still an EF1 allocation as Alice's and Carol's envy can be eliminated by ignoring their most valued item in Bob's bundle (i.e., the car) when comparing valuations.

The above allocation in our inheritance example maximizes the Nash welfare objective, which is widely considered as an important efficiency measure for allocations. It provides an appealing compromise between utilitarian and egalitarian social welfare objectives, defined as the sum of agent valuations for their allocated items and minimum valuation among all agents, respectively. Finding a Nash welfare maximizing allocation is well-known to be computationally tractable in settings with divisible items (by minimizing the convex program of Eisenberg and Gale [21]), where the resulting allocation also happens to be envy-free. In the case of indivisible goods, besides being EF1, Nash welfare maximizing allocations have some additional fairness properties (see e.g. [15]), including a remarkable *scale-invariant* property that rescaling the valuation of an agent does not change the allocation that maximizes Nash welfare (see [30] for a detailed discussion).

On the other hand, the Nash welfare maximizing allocation from the example above can hardly be considered fair, since Bob gets more value than Alice and Carol put together in our symmetric instance<sup>1</sup>. Consider another allocation, where Alice gets the ring for the value of 9, Bob gets the car for the value of 10, while the painting and the necklace go to Carol with the total value of 15. This allocation is slightly sub-optimal in terms of the Nash welfare, but it looks clearly better than the previous allocation from the fairness point of view. Indeed, observe that, inevitably, one agent must get the car; it is Bob in our case. This already creates some envy from Alice and/or Carol. So, it makes more sense to allocate the remaining three items only to them and give nothing more to Bob. More concretely, such allocations have the much more appealing property of envy-freeness up to *any* item (EFX), introduced by Caragiannis et al. [15]. Now, the envy of Alice and Bob for Carol can be eliminated by ignoring their *least-valued* item in Bob's bundle when comparing valuations. Arguably, EFX is the best fairness analog of envy-freeness for indivisible items. But even though EF1 allocations can be computed in many different ways (e.g., by a simple draft mechanism, by

<sup>1</sup>The instance is invariant under a cyclic permutation of agents' names and the respective cyclic permutation of all items except the car.

the local-search algorithm of Lipton et al. [29], by maximizing the Nash welfare, etc.), for EFX allocations, we do not even know whether they always exist. Moreover, asking for EFX allocations that simultaneously satisfy some efficiency guarantee is way beyond the reach of known techniques.

*Donation of items.* A natural way to deal with the (hypothetical) issue of the existence of EFX allocations is to consider partial allocations, which leave some of the items unallocated. The unallocated items could then be donated to a charity or sold at auction with the profit equally distributed between the agents. Actually, in situations like the above inheritance example that arise in practice, it is usually observed that people donate some of their possessions to charities before distributing the rest to their heirs. Undoubtedly, the donation of items can completely eliminate envy, e.g., by donating all items and allocating nothing to the agents. However, although beneficial for the charity cause, this outcome cannot be considered as good for the problem at hand. Therefore, an important question to ask is:

How can we achieve an EFX allocation of high Nash welfare by donating items smartly?

*Our results.* We make careful use of the idea of donating items and hit many birds with one stone. We present an algorithm that takes an allocation of maximum Nash welfare as input and, after donating some of the items, outputs an allocation of the remaining ones that is EFX and has a Nash welfare that is at least half the optimal one for the original instance (that includes the donated items). As we show, this efficiency guarantee is best possible for EFX allocations, no matter whether there are donated items or not. A remarkable bonus property is that the bundle that is allocated to each agent is a subset of the bundle he/she initially gets in the maximum Nash welfare allocation. When the same approach is applied to *large market* instances, where any item's contribution is only a small fraction of the total value of any agent, our algorithm outputs almost optimal allocations. The algorithm can be modified to work with input allocations that have sub-optimal Nash welfare. Combined with a  $\rho$ -approximation algorithm for maximizing Nash welfare (such as the algorithms of Cole and Gkatzelis [19] or Barman et al. [10]), the modified algorithm runs in polynomial-time and computes an EFX allocation that is  $2\rho$ -approximation to the maximum Nash welfare.

*Techniques.* Our algorithms are purely combinatorial. Our algorithms do not reallocate any items between the bundles in the initial allocation but permanently remove items from the instance. This removal operation gives us fine control of the agents' preferences. For example, if an agent is "happy" with a bundle from which we do not remove an item, then she stays happy with it after the removal of an item. At a high level, we always seek to find a complete EFX feasible matching of agents with bundles and carefully remove an item so that the size of the matching never decreases. From the Nash welfare maximality property, we derive that no bundle can lose more than a certain fraction of its value for the agent who had it in the initial Nash welfare maximizing allocation, yielding the desired Nash welfare guarantee. For large markets, we are able to obtain our considerably improved result by proving a number-theoretic lemma using Karamata's inequality for the logarithm function.

*Significance.* We employ the natural and practical operation of donating/removing items to resolve the challenging issues related to the EFX solution concept. Our approach contrasts with the recent attempts to mitigate such challenges by considering approximate versions of EFX (e.g., in the papers by Plaut and Roughgarden [34] and Amanatidis et al. [1]). Instead, we keep the precise definition of EFX and use approximation only for the efficiency guarantee, which is unavoidable. This approach of considering item donations could be useful in the study of the interplay of other fairness notions with efficiency.

## 1.1 Related work

*Fair division with indivisible items.* The rigorous study of fair division with divisible items has a long history; it originates from the work of Steinhaus [35] in the 40's and includes very recent breakthroughs like the envy-free cake-cutting protocol of Aziz and McKenzie [7]. In contrast, fairly allocating indivisible items among agents has not been as popular until very recently. Most probably, the reason for this delay is that the beautiful fairness notions of envy-freeness and proportionality, that have received so much attention in the literature on divisible items, are not always achievable with indivisible items. To give an embarrassing example, just consider a single item and two agents with identical valuations for the item.

The recent interest for the indivisible items setting was sparked with the definition of fairness notions that approximate envy-freeness and proportionality. In particular, the notion of maxmin fair share (MMS), defined by Budish [13] and Bouveret and Lemaître [12] can be thought of as an approximate version of proportionality and has received much attention recently, e.g., see [2, 9, 25, 27]. Besides the concepts of EF1 and EFX mentioned above, approximate versions of envy-freeness include epistemic envy-freeness [6] or notions that require the minimization of the envy-ratio [29] and degree of envy [17, 33] objectives.

As mentioned earlier, EF1 is easy to achieve with several different methods. What is really challenging is to achieve EF1 and Pareto-optimality simultaneously. This was proved to be possible in [15], by an allocation of maximum Nash welfare. The popular website Spliddit [26], available at [www.spliddit.org](http://www.spliddit.org), returns such allocations as part of its "Divide goods" application. Following [15], Barman et al. [10] investigate whether EF1 and Pareto-optimal allocations can be computed efficiently and present a pseudo-polynomial-time algorithm.

Unlike EF1, the existence of EFX allocations is still a mystery, even for three agents with additive valuations. Plaut and Roughgarden [34] prove the existence of EFX allocations for setting with two agents only or with more agents and identical valuations. In addition, they present an algorithm for computing an  $1/2$ -EFX allocation, where the value of each agent from her bundle is at least half of what the EFX property requires it to be. Caragiannis et al. [15] and Amanatidis et al. [1] consider different approximations of EFX (namely, approximations of the strongly related pairwise MMS fairness concept).

*Nash welfare.* The history of Nash social welfare (or, simply, Nash welfare) dates back to the 50's, where it was used in bargaining problems [31]. In fair division, it is considered as a good measure to balance fairness with efficiency. In particular, in settings with divisible items, maximizing the Nash welfare can be done in polynomial-time using the convex program of Eisenberg and Gale [21], which also leads to envy-free allocations. Unfortunately, such nice properties disappear in the indivisible items setting as the problem of computing an allocation of maximum Nash welfare becomes APX-hard [28] (see also [32] for a weaker hardness result). Still, as the Nash welfare is a very important efficiency measure, several constant-approximation algorithms have been proposed recently. The first such algorithm was due to Gkatzelis and Cole [19] and approximates the Nash welfare within a factor of 2.89. This was improved to 2 by Cole et al. [18], and further to 1.45 by Barman et al. [10]. See also [4, 5, 16, 24] for approximation algorithms in more general settings (with non-additive valuations). With the exception of [10] that uses item pricing techniques, rounding of convex programming relaxations is the main algorithmic tool in this line of research.

*Other related papers.* Our main positive result bounds the efficiency gap of the best EFX allocation in terms of Nash welfare. As such, this is a *price of fairness* result, a notion that was introduced independently by Bertsimas et al. [11] and Caragiannis et al. [14]. Also, the general structure of our algorithms has conceptual similarities (but also many technical differences) to algorithms for

computing combinatorial Walrasian equilibria [22], where items are packaged into groups instead of being donated. Finally, we remark that large market assumptions have been considered in a few different areas recently, such as in the study of the price of anarchy of large games [20, 23], in budget-feasible mechanisms [3], and in optimizing over (very different from ours) fairness constraints [8].

## 1.2 Roadmap

The rest of the paper is structured as follows. We begin with formal definitions in Section 2. Our counter-example showing that the Nash welfare of EFX allocations can be far from optimal is given in Section 3. The description and analysis of our main positive result is presented in Section 4. Our modifications that allow the algorithm to work with allocations of sub-optimal Nash welfare as input are given in Section 5. The large market assumption is considered in Section 6. We conclude in Section 7 with a short discussion and open problems.

## 2 PRELIMINARIES

Let  $\mathcal{N}$  be the set of  $n$  agents and  $\mathcal{M}$  be the set of  $m$  indivisible items. We enumerate agents and items from 1 through  $n$  and through  $m$ , respectively. We will often refer to the set of agents  $\mathcal{N}$  as  $[n]$  and to the set of items  $\mathcal{M}$  as  $[m]$ , where  $[k] \stackrel{\text{def}}{=} \{1, 2, \dots, k\}$  for any positive integer number  $k$ . Each agent  $i \in \mathcal{N}$  is endowed with a valuation function  $v_i : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$ , where  $\mathbb{R}_{\geq 0}$  is the set of non-negative real numbers. Valuations are additive; by slight abuse of notation, agent  $i$  has a value of  $v_i(S) \stackrel{\text{def}}{=} \sum_{g \in S} v_i(g)$  for any set  $S \subseteq \mathcal{M}$ .

We refer to a set of items as a *bundle* and extensively consider partitions of items into bundles. Formally, a *partition* of a set of items  $S \subseteq \mathcal{M}$  is a set of  $n$  disjoint bundles  $\mathbf{B} = \{B_1, \dots, B_n\}$  such that  $B_i \cap B_j = \emptyset$  and  $\bigcup_{i \in [n]} B_i = S$ . An *allocation*  $\mathbf{X} = (X_1, \dots, X_n)$  of the items in  $S \subseteq \mathcal{M}$  to the agents of  $\mathcal{N}$  is an ordered partition of the item set  $S$ , where each agent  $i \in [n]$  gets the bundle  $X_i$ . We denote the set of all possible allocations of  $S$  to the agents of  $\mathcal{N}$  as  $\mathcal{X}(S)$ . From now on, the term allocation implies an allocation to the agents of  $\mathcal{N}$ .

An allocation  $\mathbf{X} = (X_1, \dots, X_n)$  is called

- *envy-free (EF)* if  $v_i(X_i) \geq v_i(X_j)$  for any  $i, j \in \mathcal{N}$ ;
- *envy-free up to one item (EF1)* if  $v_i(X_i) \geq \min_{g \in X_j} v_i(X_j - \{g\})$  for any  $i, j \in \mathcal{N}$ ;
- *envy-free up to any item (EFX)*<sup>2</sup> if  $v_i(X_i) \geq \max_{g \in X_j} v_i(X_j - \{g\})$  for any  $i, j \in \mathcal{N}$ .

We measure the efficiency of an allocation  $\mathbf{X} = (X_1, \dots, X_n)$  using its *Nash social welfare* (or, simply *Nash welfare*)  $\text{NW}(\mathbf{X})$ , which is defined as the geometric mean of  $\{v_i(X_i)\}_{i=1}^n$ . For any set of items  $S \subseteq \mathcal{M}$ , we denote by  $\text{opt}(S)$  the maximum Nash welfare over all allocations of  $S$ . Formally,

$$\text{NW}(\mathbf{X}) \stackrel{\text{def}}{=} \left( \prod_{i \in [n]} v_i(X_i) \right)^{1/n}, \quad \text{opt}(S) \stackrel{\text{def}}{=} \max_{\mathbf{X} \in \mathcal{X}(S)} \text{NW}(\mathbf{X}).$$

Our goal is to find an EFX allocation with as large Nash welfare as possible. We say that the allocation  $\mathbf{X}$  of  $S \subseteq \mathcal{M}$  is  $\alpha$ -efficient if  $\alpha \cdot \text{NW}(\mathbf{X}) \geq \text{opt}(S)$ . An allocation  $\mathbf{X}$  is called *Pareto-optimal* if there is no other allocation  $\mathbf{Y}$  with at least as high value for every agent  $i$ , i.e.,  $v_i(Y_i) \geq v_i(X_i)$ , and strictly higher value for one of the agents. In particular, a Nash welfare maximizing allocation is Pareto-optimal.

<sup>2</sup>If the set  $X_j$  is empty we assume that  $\min_{g \in X_j} v_i(X_j - \{g\}) = \max_{g \in X_j} v_i(X_j - \{g\}) = 0$ . We remark that our definition of EFX is slightly different than the one in [15], where the maximum is taken over the items of bundle  $X_j$  for which agent  $i$  has strictly positive valuation. However, our positive results are only stronger in this way, and in the proof of Theorem 1, all agent valuations are non-zero.

We now introduce *large markets*. These are allocation problems, in which the valuation of every agent for any item is a small fraction of her total value for all items of  $\mathcal{M}$ .

**DEFINITION 1 (LARGE MARKET).** *An allocation problem satisfies the large market condition with parameter  $\varepsilon \in (0, 1]$  if  $v_i(g) \leq \frac{\varepsilon}{n} v_i(\mathcal{M})$  for every agent  $i \in [n]$  and any item  $g \in \mathcal{M}$ .*

In Section 6, we actually use the following weaker large market condition, which is defined with respect to any allocation of optimal Nash welfare.

**DEFINITION 2 (LARGE MARKET W.R.T. OPTIMAL ALLOCATION X).** *An allocation X of optimal Nash welfare satisfies the large market condition with parameter  $\varepsilon \in (0, 1]$  if  $v_i(g) \leq \varepsilon \cdot v_i(X_i)$  for every agent  $i \in [n]$  and any item  $g \in X_i$ .*

**CLAIM 1.** *The large market condition with parameter  $\varepsilon$  implies the large market condition for an optimal Nash welfare allocation with parameter  $\varepsilon' = \varepsilon / (1 - \frac{n-1}{n}\varepsilon)$ .*

**PROOF.** We use the fact that a Nash welfare maximizing allocation X is also EF1 [15]. Let us fix any agent  $i \in \mathcal{N}$ . The EF1 condition for agent  $i$  implies that  $\forall j \neq i, \exists g_{ij} \in X_j$  such that  $v_i(X_i) \geq v_i(X_j \setminus \{g_{ij}\})$ . By summing these inequalities over all  $j \neq i$  and the equality  $v_i(X_i) = v_i(X_i)$ , we get

$$n \cdot v_i(X_i) \geq v_i(X_i) + \sum_{j \neq i} v_i(X_j \setminus \{g_{ij}\}) = v_i(\mathcal{M}) - \sum_{j \neq i} v_i(g_{ij}) \geq v_i(\mathcal{M}) - (n-1) \cdot \frac{\varepsilon}{n} v_i(\mathcal{M}).$$

Therefore, for any agent  $i$  and item  $g \in X_i$  we have

$$v_i(g) \leq \frac{\varepsilon}{n} v_i(\mathcal{M}) \leq \frac{\varepsilon}{n} n \cdot v_i(X_i) \left(1 - \frac{n-1}{n}\varepsilon\right)^{-1} = \varepsilon' \cdot v_i(X_i). \quad \square$$

### 3 A LOWER BOUND ON THE NASH WELFARE OF EFX ALLOCATIONS

Before presenting our positive results, we give an allocation problem where almost half of the optimal Nash welfare must be sacrificed in order to achieve EFX, either for the original set of items or for any subset of them.

**THEOREM 1.** *For any positive integer  $n$  and  $\varepsilon > 0$ , there is an allocation problem with  $n$  agents and set of items  $\mathcal{M}$ , such that any EFX allocation X of any subset of items  $S \subseteq \mathcal{M}$  satisfies  $(2^{1-1/n} - 2\varepsilon) \cdot \text{NW}(\mathbf{X}) \leq \text{opt}(\mathcal{M})$ .*

**PROOF.** Consider the allocation problem with  $n$  agents, set of items  $\mathcal{M} \stackrel{\text{def}}{=} \{1, \dots, 2n-1\}$ , and valuation function  $v_i : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$  of agent  $i \in [n]$  that is defined as follows:

$$v_i(g) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{when } g \in \{1, \dots, n-1\}, \\ 1 - \varepsilon & \text{when } g = 2n - i, \\ \varepsilon/2n, & \text{otherwise.} \end{cases}$$

The allocation  $\mathbf{X}^* = (X_1^*, \dots, X_n^*)$  with  $X_i^* = \{i, 2n - i\}$  for  $i \in [n-1]$ , and  $X_n^* = \{n\}$  has Nash welfare  $\text{NW}(\mathbf{X}^*) = \sqrt[n]{(2 - \varepsilon)^{n-1}(1 - \varepsilon)}$ . Hence, the optimal Nash welfare is

$$\begin{aligned} \text{opt}(\mathcal{M}) &\geq \sqrt[n]{(2 - \varepsilon)^{n-1}(1 - \varepsilon)} = 2^{1-1/n} \cdot \left(1 - \frac{\varepsilon}{2}\right)^{1-1/n} (1 - \varepsilon)^{1/n} \\ &> 2^{1-1/n} \cdot (1 - \varepsilon)^{1-1/n} (1 - \varepsilon)^{1/n} > 2^{1-1/n} - 2\varepsilon. \end{aligned}$$

Now consider any EFX allocation  $\mathbf{X} = (X_1, \dots, X_n)$  of a set of items  $S \subseteq \mathcal{M}$ . By the pigeonhole principle, there must be some agent  $j$  who does not receive any of the first  $n-1$  items. Then, agent  $j$ 's value is at most  $v_j(X_j) \leq \sum_{g=n}^{2n-1} v_i(g) < 1 - \frac{\varepsilon}{2}$ . It means that, if any other agent  $i \neq j$  receives

a large item  $g \in [n - 1]$ , then her allocation must contain only that item, i.e.,  $X_i = \{g\}$ . Indeed, if there is an item  $g'$  different than  $g$  so that  $g$  and  $g'$  belong to bundle  $X_i$ , then the EFX condition is violated, as  $v_j(X_j) < 1 - \varepsilon/2 < 1 \leq v_j(X_i \setminus g')$ . This means that no agent gets value higher than 1: either agent  $i$  gets a large item  $g \in [n - 1]$  and, subsequently,  $v_i(X_i) = v_i(g) = 1$ , or she gets (a subset of) items  $n, n + 1, \dots, 2n - 1$  and has value  $v_i(X_i) \leq \sum_{g=n}^{2n-1} v_i(g) < 1 - \varepsilon/2$ . Therefore,  $\text{NW}(\mathbf{X}) \leq 1$  which, together with the inequality on  $\text{opt}(\mathcal{M})$  above, completes the proof.  $\square$

#### 4 MAIN ALGORITHM

We now present our algorithm for computing an EFX allocation of some of the items to the agents. Together with the allocation problem (a set of items  $\mathcal{M}$  and  $n$  agents with valuations for the items in  $\mathcal{M}$ ), the algorithm receives as input a Nash welfare maximizing allocation  $\mathbf{X}$  for it. It tries to match as many bundles from the initial allocation to agents as possible and repeatedly removes items from the bundles as long as this matching does not correspond to an EFX allocation. The algorithm is guaranteed to output an EFX allocation  $\mathbf{Y}$  with Nash welfare which is at least half of the original one. More precisely, we will show that  $2^{\frac{n-1}{n}} \cdot \text{NW}(\mathbf{Y}) \geq \text{NW}(\mathbf{X}) = \text{opt}(\mathcal{M})$ .

Before presenting our algorithm in detail, we introduce several useful notions, giving forward pointers to the lines of the pseudocode (see Algorithm 1) where these notions are used. During its execution, the algorithm maintains a disjoint set  $\mathbf{Z}$  of  $n$  bundles. Initially (see Line 1),  $\mathbf{Z}$  consists of the bundles in  $\mathbf{X}$ . The algorithm progresses in rounds. In each round (defined by an execution of Lines 3-12 of the pseudocode in the “repeat-until” loop), it tries to compute an EFX allocation (with particular properties) by assigning the bundles of  $\mathbf{Z}$  to the agents. Whenever this is not possible, it removes an item from a bundle of  $\mathbf{Z}$  and proceeds to the next round. When a bundle of  $\mathbf{Z}$  misses an item, it becomes a *touched* bundle.

In order to compute the EFX allocation in each round, the algorithm uses the notion of the *EFX feasibility graph*, which is defined as follows. The EFX feasibility graph is a bipartite graph  $G = (\mathcal{N} \cup \mathbf{Z}, E)$  between two sets of vertices: the first part contains all agents  $\mathcal{N}$ , each agent as a vertex; the second part contains the bundles in  $\mathbf{Z}$ , each bundle  $Z_i$  as a vertex. Edges of  $G$  are defined as

$$E(G) \stackrel{\text{def}}{=} \{(i, Z_j) \mid \text{(i) } Z_j \text{ is EFX feasible for } i; \text{ (ii) } v_i(Z_j) > v_i(Z_i) \text{ if } i \neq j\}.$$

In particular, EFX feasibility in condition (i) requires that  $v_i(Z_j) \geq \max_{g \in Z_k} v_i(Z_k - \{g\})$  for all  $k \in [n]$ . Condition (ii) in the definition of  $E(G)$  restricts the space of possible allocations and expresses our preference for matching agents to their initial bundles whenever possible. The call of the subroutine `EFX-feasibility-graph` in Line 3 of Algorithm 1 builds the EFX feasibility graph for the current set of bundles  $\mathbf{Z}$ .

In each round, the algorithm computes a matching in the EFX feasibility graph. This matching can naturally be thought of as a partial allocation of the current bundles to the agents. If the matching is perfect, the corresponding (complete) allocation is returned and the algorithm terminates. Otherwise, an item is removed from some bundle and the algorithm proceeds to the next round. The matching computed in each round has particular properties that guarantee that the algorithm makes progress during its execution. In particular, all touched bundles are matched (condition (a) in Line 6). Under this condition, the matching contains the maximum number of edges of the form  $(i, Z_i)$  (condition (b) in Line 7). And, under these two conditions, the size of the matching is maximized (condition (c) in Line 8).

The algorithm finds the *robust demand* bundle of an arbitrary unmatched agent  $i$  (Line 10) if the current matching is not perfect. The robust demand for agent  $i$  is defined as

$$\text{robust-demand}(i, \mathbf{Z}) \stackrel{\text{def}}{=} Z_j \in \underset{Z \in \mathbf{Z}}{\text{argmax}} \{ \max_{c \in Z} v_i(Z \setminus \{c\}) \},$$

breaking ties arbitrarily. Then, the algorithm updates the set of bundles  $Z$  by removing from the robust-demand bundle  $Z_j$  the least valued item of agent  $i$  (Line 11). The definition of the robust demand guarantees that the bundle  $Z_j$  will be EFX feasible for agent  $i$  in the next round. In addition, as we will see, the edge  $(i, Z_j)$  will belong to the EFX feasibility graph in the next iteration.

**Input:** allocation  $X = (X_1, \dots, X_n)$  of  $\mathcal{M}$ , such that  $\text{NW}(X) = \text{opt}(\mathcal{M})$ .  
**Output:** allocation  $Y = (Y_1, \dots, Y_n)$  of a set  $S \subseteq \mathcal{M}$ .

- 1 Let  $Z = (Z_1, \dots, Z_n) \leftarrow (X_1, \dots, X_n)$  be an ordered partition of  $\mathcal{M}$ .
- 2 **repeat**
- 3      $G \leftarrow \text{EFX-feasibility-graph}([n], Z)$ ;
- 4     Let  $T$  be the set of touched bundles in  $Z$ ;
- 5     Let  $M$  be a matching in  $G$  such that
- 6         (a) All bundles in  $T$  are matched in  $M$ ,
- 7         (b) Under (a),  $|M \cap \{(i, Z_i) : i \in [n]\}|$  is maximized, and
- 8         (c) Under (a) and (b),  $|M|$  is maximized;
- 9     **if**  $\exists i \in [n]$  not matched in  $M$  **then**
- 10          $Z_j \leftarrow \text{robust-demand}(i, Z)$ ;
- 11          $Z \leftarrow (Z_1, \dots, Z_{j-1}, Z_j \setminus \{c\}, Z_{j+1}, \dots, Z_n)$ , where  $c \in \{\text{argmin}_{g \in Z_j} v_i(g)\}$ ;
- 12     **end**
- 13 **until**  $|M| = n$ ;
- 14 **return**  $Y = (M_1, \dots, M_n)$ ; ▷ It must hold  $\{M_i = Z_i\}_{i \in [n]}$ .

**Algorithm 1:** Computes an EFX allocation of high Nash welfare.

Algorithm 1 will eventually terminate and output some allocation  $Y$ , since in every iteration of the repeat loop (except the last time  $\bar{t}$ ) it removes an item from  $\mathcal{M}$ . Next, we need to argue that in every iteration  $t \in \{0, \dots, \bar{t}\}$  of the repeat loop the algorithm is correct, i.e., that in Line 6 of Algorithm 1, all touched bundles from  $T(t)$  can be simultaneously matched in graph  $G(t) \stackrel{\text{def}}{=} \text{EFX-feasibility-graph}([n], Z(t))$ <sup>3</sup>. This is proved in Lemma 1; the proof uses several useful observations about the structure of the EFX feasibility graph  $G(t)$ .

**LEMMA 1.** *All bundles in  $T(t)$  can be simultaneously matched to agents in  $G(t)$  at any iteration  $t$ .*

**PROOF.** The proof proceeds by induction on  $t$ . For  $t = 0$ , the statement is true, as  $T(0) = \emptyset$  and we can choose  $M(0) = \emptyset$ . We assume that at time  $t < \bar{t}$  the matching  $M(t)$  covers all bundles from  $T(t)$ . Let us assume that agent  $i^*$  is chosen in the execution of the if statement (Line 9 of Algorithm 1) at time  $t$  and the item  $c^* \in \{\text{argmin}_{g \in Z_{j^*}(t)} v_{i^*}(g)\}$  is removed from the bundle  $Z_{j^*}(t) = \text{robust-demand}(i^*, Z(t))$  in Lines 10-11. Now, in order to show that all bundles in  $T(t+1)$  can be matched to agents in  $G(t+1)$ , we will need a few useful observations about the edges of  $G(t+1)$ .

**CLAIM 2.** *Graph  $G(t+1)$  contains every edge  $e = (i, Z_j)$  with  $Z_j \neq Z_{j^*}$  of the graph  $G(t)$ .*

**PROOF.** Indeed, since no item is removed from bundle  $Z_j$ , we have  $v_i(Z_j(t+1)) = v_i(Z_j(t)) \geq v_i(Z_i(t)) \geq v_i(Z_i(t+1))$ . Similarly,  $Z_j(t+1)$  remains EFX feasible for agent  $i$ . Thus  $(i, Z_j) \in E(G(t+1))$ .  $\square$

**CLAIM 3.** *There is no edge between  $i^*$  and  $Z_{i^*}$  in  $G(t)$ .*

<sup>3</sup>In the  $t^{\text{th}}$  iteration,  $Z(t)$  and  $T(t)$  represents  $Z$  and  $T$  before the execution of the if statement in Lines 9-12 of the algorithm.

PROOF. We prove this claim by contradiction. Suppose the edge  $(i^*, Z_{i^*})$  belongs to  $E(G(t))$ , but  $i^*$  was not matched in  $M(t)$ . Then we may increase the number of matched pairs  $(i, Z_i)$  (condition (b)) in  $M(t)$  by adding or possibly substituting another edge to  $Z_{i^*}$  in  $M(t)$ . As the set of matched bundles in  $Z(t)$  can only increase after such an operation, we get a contradiction to condition (b) (Line 7 of Algorithm 1).  $\square$

CLAIM 4. *There is an edge between  $i^*$  and  $Z_{j^*}$  in both graphs  $G(t)$  and  $G(t + 1)$ .*

PROOF. Claim 3 says that  $Z_{i^*}(t)$  is not EFX feasible for agent  $i^*$ , i.e., there exists agent  $i \in [n]$  such that  $v_{i^*}(Z_{i^*}(t)) < v_{i^*}(Z_i(t) \setminus \{g\})$  for some item  $g \in Z_i(t)$ . As bundle  $Z_{j^*}(t)$  is the robust demand of  $i^*$  at time  $t$ , we have

$$v_{i^*}(Z_{j^*}(t + 1)) = v_{i^*}(Z_{j^*}(t) \setminus \{c^*\}) \geq \max_{c \in Z_i(t)} v_{i^*}(Z_i(t) \setminus \{c\}) > v_{i^*}(Z_{i^*}(t)) \geq v_{i^*}(Z_{i^*}(t + 1)).$$

Similarly,  $Z_{j^*}(t)$  and  $Z_{j^*}(t + 1)$  are EFX feasible for agent  $i^*$ , as  $v_{i^*}(Z_{j^*}(t) \setminus \{c^*\})$  is at least as high as  $v_{i^*}(Z_j(t) \setminus \{g\})$  for all  $j \in [n]$  and any item  $g \in Z_j(t)$ . Therefore,  $(i^*, Z_{j^*}) \in E(G(t + 1))$  and  $(i^*, Z_{j^*}) \in E(G(t))$ .  $\square$

We can now show that the algorithm will match  $Z_{j^*}$  to some agent.

CLAIM 5. *Bundle  $Z_{j^*}$  is matched in  $M(t)$  to an agent  $k$ .*

PROOF. If the bundle  $Z_{j^*}$  is unmatched in  $M(t)$ , then there is a bigger matching  $M(t) \cup (i^*, Z_{j^*})$  in  $G(t)$ : by Claim 4, the edge  $(i^*, Z_{j^*})$  belongs to  $E(G(t))$  and, as the agent chosen when Line 9 is executed during iteration  $t$ ,  $i^*$  is unmatched in  $M(t)$ . This contradicts property (c) in Line 8 of Algorithm 1.  $\square$

Finally, we note that  $T(t + 1) = T(t) \cup Z_{j^*}$ , as bundle  $Z_{j^*}$  is the only new bundle that can become touched in iteration  $t$ . Now consider the matching

$$M' \stackrel{\text{def}}{=} (M(t) \setminus \{(k, Z_{j^*})\}) \cup \{(i^*, Z_{j^*})\} \quad (1)$$

in  $G(t + 1)$  that covers all bundles in  $T(t + 1)$ . Indeed, according to Claim 2,  $G(t + 1)$  contains all edges of  $M(t) \setminus \{(k, Z_{j^*})\}$  and, according to Claim 4, the edge  $(i^*, Z_{j^*})$  belongs to  $G(t + 1)$  as well. Hence,  $M'$  is a matching in  $G(t + 1)$ . Since (by our induction hypothesis) all bundles of  $T(t) \setminus \{Z_{j^*}\}$  are matched in  $M(t) \setminus \{(k, Z_{j^*})\}$ ,  $M'$  covers all bundles in  $T(t + 1) = T(t) \cup Z_{j^*}$ . The proof of the induction step is complete.  $\square$

As the algorithm terminates after finding a complete matching in the EFX feasibility graph, the returned solution  $Y$  must be an EFX allocation. We note that Algorithm 1 runs in polynomial time: it executes at most  $m$  iterations and all steps in each iteration can be completed in polynomial time. Indeed, the only non trivial part is the computation of matching  $M$  (Line 5) under the conditions (a)-(c) (Lines 6-8). To this end, we can assign weights to the edges of the graph  $G$  that express our preferences (a)-(c) and compute a maximum weighted matching in the weighed version of  $G$ .<sup>4</sup> For example, we can give a weight of  $n^4$  to the edges of  $G$  incident to bundle  $Z_i \in T$ , and then add an extra weight of  $n^2$  to edges  $(i, Z_i) \in E(G)$  for  $i \in [n]$ , and give a weight of 1 to the remaining edges of  $G$ .

In the following we give efficiency guarantees for the returned EFX allocation  $Y$ . Very informally, Algorithm 1 does not remove too many items. First, we observe that at the end of the algorithm there must be at least one untouched bundle, i.e., a bundle from which no items have been removed.

CLAIM 6. *There is an untouched bundle  $Z_i = X_i$  upon termination of Algorithm 1.*

<sup>4</sup>We remark that this is apparently not the fastest way to compute the desired matching.

PROOF. Assume to the contrary that the algorithm has removed items from every bundle in  $Z$  after iteration  $t$ , i.e., the set of touched bundles becomes the set  $T(t+1) = \{Z_i\}_{i \in [n]}$  of size  $n$  after iteration  $t$ . Then, by Lemma 1, the algorithm will match all bundles in  $M(t+1)$  and terminate at round  $t+1$ . Let agent  $i^*$  and bundle  $Z_{j^*}$  be the ones chosen, respectively, when Algorithm 1 executes Lines 9 and 10 at iteration  $t$ . At this time, we have  $Z_j \in T(t)$  for all  $j \neq j^*$ . Thus, all touched bundles  $Z_j$  with  $j \neq j^*$  must be matched in  $M(t)$  (due to the condition (a) in Line 6). According to Claim 5, bundle  $Z_{j^*}$  must also be matched in  $M(t)$ . Therefore,  $|M(t)| = n$  and the algorithm should have terminated after iteration  $t$  leaving bundle  $Z_{j^*}$  untouched.  $\square$

We are now ready to present guarantees for the Nash welfare of allocation  $Y$ .

LEMMA 2.  $v_i(X_i) \leq 2 \cdot v_i(Y_i)$  for any agent  $i \in [n]$  and there exists an agent  $i^o \in [n]$  such that  $v_{i^o}(X_{i^o}) \leq v_{i^o}(Y_{i^o})$ .

PROOF. We note first that when the algorithm terminates at time  $\bar{t}$ , then  $v_i(Z_i) \leq v_i(Y_i)$  for all  $i \in [n]$ . Indeed, this is the case if agent  $i$  is matched to bundle  $Z_i$ . If  $i$  is matched to another bundle  $Y_i = Z_j$  in  $G(\bar{t})$ , then  $v_i(Z_j) > v_i(Z_i)$  by the definition of the EFX feasibility graph  $G(\bar{t})$ . Thus, Claim 6 proves the second part of the lemma, since it states that there exists some agent  $i^o$  such that  $v_{i^o}(X_{i^o}) = v_{i^o}(Z_{i^o}) \leq v_{i^o}(Y_{i^o})$ .

To complete the proof of the lemma, it is sufficient to show that  $2 \cdot v_i(Z_i) \geq v_i(X_i)$  for any agent  $i \in [n]$ . We do so by contradiction. If this is not the case, let  $t \in \{0, \dots, \bar{t}\}$  be the first time when  $2 \cdot v_i(Z_i(t))$  becomes strictly smaller than  $v_i(X_i)$  for an agent  $i$  after an item was removed. Let  $i^*$  be the agent who was chosen in Line 9 and  $Z_{j^*}$  be the bundle chosen in Line 10 of the algorithm at iteration  $t$ . Then  $j^*$  is the agent for whom  $2 \cdot v_{j^*}(Z_{j^*}(t))$  became smaller than  $v_{j^*}(X_{j^*})$ . For convenience of notation, we denote the matching  $\{(i, Z_i)\}_{i \in [n]}$  by  $M^o$ . We consider the matching  $M' \triangleq M(t) \setminus \{(k, Z_{j^*})\} \cup \{(i^*, Z_{j^*})\}$ , that was also used in the proof of Lemma 1. Recall that, by Claim 5, bundle  $Z_{j^*}$  is matched to agent  $k$  in  $M(t)$ . As agent  $k$  is not matched in  $M'$ , we have  $|M'| < n = |M^o|$ . Therefore, one can represent the union of the matchings  $M'$  and  $M^o$  as a collection of augmenting paths and cycles, including degenerate cycles that consist of a single edge that belongs to both  $M'$  and  $M^o$ .

First, assume that  $M' \cup M^o$  has a non-degenerate cycle  $C$  with  $2k$  vertices  $\{j_1, \dots, j_k\}$  and  $\{Z_{j_1}, \dots, Z_{j_k}\}$ .

$$C \stackrel{\text{def}}{=} \{(j_1, Z_{j_2}), \dots, (j_{k-1}, Z_{j_k}), (j_k, Z_{j_1})\} \in M', \quad \{(j_1, Z_{j_1}), \dots, (j_k, Z_{j_k})\} \in M^o.$$

For convenience of notation, let us denote  $j_{k+1} = j_1$ . Since each  $(j_i, Z_{j_{i+1}}) \in M'$  is an edge in the EFX feasibility graph  $G(t)$ , we have  $v_{j_i}(Z_{j_{i+1}}) > v_{j_i}(Z_{j_i})$  for every  $i \in [k]$ . This implies that  $v_{j_i}(X_{j_i}) + v_{j_i}(Z_{j_{i+1}}) - v_{j_i}(Z_{j_i}) > v_{j_i}(X_{j_i})$  for every  $i \in [k]$ . Thus, we get a Pareto improvement in the allocation  $X$  for each agent  $j_i \in C$  by replacing bundle  $X_{j_i}$  with bundle  $(X_{j_i} \setminus Z_{j_i}) \cup Z_{j_{i+1}}$  for every  $i \in [k]$ . Hence, we can improve the Nash welfare of the allocation  $X$ , a contradiction to its optimality.

Second, we observe that bundle  $Z_{j^*}$  is matched in  $M'$  to  $i^*$  which, by Claim 3, must be different from  $j^*$ . Since  $M' \cup M^o$  does not have non-degenerate cycles and bundle  $Z_{j^*}$  is not matched to  $j^*$  in  $M'$ , bundle  $Z_{j^*}$  must belong to an augmenting path  $P$  that originates from an unmatched bundle  $Z_{j_1}$  in  $M'$ ; notice that  $Z_{j_1}$  is unmatched in  $M(t)$  as well, since matching  $M(t)$  has the same set of bundles as  $M'$ . The augmenting path  $P$  consists of  $k$  agent vertices  $j_1, \dots, j_k$  and  $k+1$  bundle vertices  $Z_{j_1}, \dots, Z_{j_{k+1}}$ , where  $j_{k+1} = j^*$ , and  $2k$  edges:

$$P \stackrel{\text{def}}{=} \{(j_1, Z_{j_2}), \dots, (j_k, Z_{j_{k+1}})\} \in M', \quad \{(j_1, Z_{j_1}), \dots, (j_k, Z_{j_k})\} \in M^o. \quad (2)$$

We consider the following transformation  $\widehat{\mathbf{X}}$  of the initial allocation  $\mathbf{X}$  (see Figure 1):

$$\widehat{\mathbf{X}} \stackrel{\text{def}}{=} \begin{cases} \widehat{X}_{j_1} = X_{j_1} \cup Z_{j_2}, \\ \widehat{X}_{j_i} = X_{j_i} \setminus Z_{j_i} \cup Z_{j_{i+1}}, & i \in \{2, \dots, k\} \\ \widehat{X}_{j_{k+1}} = X_{j_{k+1}} \setminus Z_{j_{k+1}} \\ \widehat{X}_j = X_j, & j \notin \{j_1, \dots, j_{k+1}\} \end{cases}$$

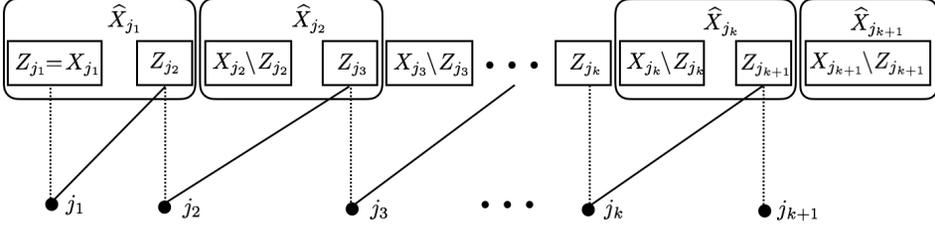


Fig. 1. The main argument in the proof of Lemma 2: reallocating the items of allocation  $\mathbf{X}$  to obtain another allocation  $\widehat{\mathbf{X}}$  with improved Nash welfare.

Let us denote the set of agents  $\{j_1, \dots, j_{k+1}\}$  as  $J$ . Similar to the previous case where we considered a non-degenerate cycle, we have  $v_{j_i}(Z_{j_{i+1}}) > v_{j_i}(Z_{j_i})$  for every  $i \in [k]$ . By our assumption for agent  $j^*$  (with  $j^* = j_{k+1}$ ), it holds that  $v_{j_{k+1}}(Z_{j_{k+1}}) < \frac{1}{2} \cdot v_{j_{k+1}}(X_{j_{k+1}})$ . Finally, since  $Z_{j_1}$  is unmatched in  $M(t)$  it must also be an untouched bundle, i.e.,  $Z_{j_1} = X_{j_1}$ . Combining these properties, we get:

$$\begin{aligned} \text{NW}(\widehat{\mathbf{X}})^n &= v_{j_1}(X_{j_1} \cup Z_{j_2}) \cdot v_{j_{k+1}}(X_{j_{k+1}} \setminus Z_{j_{k+1}}) \cdot \prod_{i=2}^k v_{j_i}(X_{j_i} \setminus Z_{j_i} \cup Z_{j_{i+1}}) \cdot \prod_{j \notin J} v_j(X_j) \\ &= [v_{j_1}(X_{j_1}) + v_{j_1}(Z_{j_2})] \cdot [v_{j_{k+1}}(X_{j_{k+1}}) - v_{j_{k+1}}(Z_{j_{k+1}})] \\ &\quad \cdot \prod_{i=2}^k [v_{j_i}(X_{j_i}) - v_{j_i}(Z_{j_i}) + v_{j_i}(Z_{j_{i+1}})] \cdot \prod_{j \notin J} v_j(X_j) \\ &> [v_{j_1}(X_{j_1}) + v_{j_1}(Z_{j_1})] \cdot \frac{1}{2} v_{j_{k+1}}(X_{j_{k+1}}) \cdot \prod_{i=2}^k v_{j_i}(X_{j_i}) \cdot \prod_{j \notin J} v_j(X_j) = \text{NW}(\mathbf{X})^n \end{aligned}$$

We have reached a contradiction, since  $\text{NW}(\mathbf{X}) = \text{opt}(\mathcal{M})$ .  $\square$

Using Lemma 1 and 2, we obtain the following statement.

**THEOREM 2.** *Given a Nash welfare maximizing allocation  $\mathbf{X}$ , Algorithm 1 computes in polynomial time a  $2^{1-\frac{1}{n}}$ -efficient EFX and Pareto-optimal allocation  $\mathbf{Y}$  of  $\cup_{i=1}^n Y_i$ , such that  $Y_i \subseteq X_i$  for all  $i \in \mathcal{N}$ .*

**PROOF.** The correctness of the algorithm was shown in Lemma 1. According to the definition of the EFX feasibility graph  $G$ , the final allocation, which is a complete matching in  $G$ , must be an EFX allocation. Finally, according to Lemma 2, the Nash welfare of the final allocation  $\mathbf{Y}$  is at least a  $2^{-(n-1)/n}$ -fraction of  $\text{opt}(\mathcal{M}) = \text{NW}(\mathbf{X})$ .

Moreover, if the returned complete matching  $M$  is different from the matching  $M^o \triangleq \{(i, Z_i)\}_{i \in [n]}$ , then there must be a cycle  $C : \{(j_1, Z_{j_2}), \dots, (j_{k-1}, Z_{j_k}), (j_k, Z_{j_1})\} \in M, \{(j_1, Z_{j_1}), \dots, (j_k, Z_{j_k})\} \in M^o$  in  $M \cup M^o$ . Then one can get a strict improvement to the initial Nash welfare maximizing

allocation  $\mathbf{X}$ , by setting  $\widehat{X}_{j_i} = X_{j_i} \setminus Z_{j_i} \cup Z_{j_{i+1}}$ , where  $j_{k+1} \stackrel{\text{def}}{=} j_1$ . Indeed,  $v_{j_i}(Z_{j_{i+1}}) > v_{j_i}(Z_{j_i})$  and  $v_{j_i}(\widehat{X}_{j_i}) > v_{j_i}(X_{j_i})$  for all  $i \in [k]$ , as  $(j_i, Z_{j_{i+1}})$  is an edge in the EFX feasibility graph  $G$ . Hence,  $M(i) = Z_i$  for all  $i \in [n]$  and  $Y_i \subseteq X_i$ .

Finally, allocation  $\mathbf{Y}$  is Pareto optimal, as otherwise one could first Pareto improve the allocation  $\mathbf{Y}$  and then give back to each agent  $i$  their original removed items  $X_i \setminus Y_i$ ; the resulting allocation would be a Pareto-improvement of the initial allocation  $\mathbf{X}$ , contradicting its optimality.  $\square$

## 5 SUBOPTIMAL INPUT ALLOCATION

One might hope that the algorithm we presented in Section 4 could also work with an input allocation of suboptimal Nash welfare. Indeed, the algorithm would result in an EFX allocation in this case too. Unfortunately, the proof of the efficiency guarantee in Lemma 2 crucially relies on the optimality of the input allocation. In this section we present a modified Algorithm 2 with a good efficiency guarantee, provided that the input allocation is efficient as well, albeit not necessarily optimal. In particular, starting with an initial allocation  $\mathbf{X}$ , the modified algorithm computes either an EFX allocation  $\mathbf{Y}$  of some of the items in  $\mathcal{M}$  that recovers a large fraction of the Nash welfare of the initial allocation or indicates that the Nash welfare of the input allocation can be improved significantly and provides such an improved allocation  $\widehat{\mathbf{X}}$  as output. Repeating the algorithm with the allocation  $\widehat{\mathbf{X}}$  as input until it produces an EFX allocation will yield, in polynomial time, an EFX allocation with at least half the Nash welfare of the initial input allocation.

The general structure of the algorithm is the same as before. It proceeds in rounds (defined by the outer repeat-until loop; Lines 5-30). In each round, it tries to match as many bundles from the initial allocation to agents as possible (by essentially repeating the matching computation on the EFX feasibility graph; Lines 5-10) and repeatedly removes items from the bundles as long as this matching does not correspond to an EFX allocation. The main difference of the modified algorithm is in the selection of the agent and an item in their robust demand bundle to be removed at each round, which is implemented in lines 11-30 and includes the inner repeat-until loop. This step is more complicated now and may modify the computed matching as well (but clearly takes polynomial time). We exploit paths on the union of matching  $M$  with the identity matching  $M^o \triangleq \{(i, Z_i) : i \in [n]\}$ . In particular, given an unmatched bundle  $Z_{j_i}$  (defined in line 11), the algorithm first computes the augmenting path of  $M \cup M^o$  that originates from vertex  $Z_{j_i}$  of the EFX feasibility graph. This is a path that alternates between edges of matching  $M^o$  and  $M$ , and terminates with an edge of  $M^o$  and the vertex corresponding to the unmatched agent  $j_k$ . The algorithm temporarily selects agent  $j_k$  (Line 14, where the set of edges of  $M$  that belong to the augmenting path is returned by the call to the subroutine `augmenting-path`) and computes agent's  $j_k$  robust demand (Line 15), say bundle  $Z_{j^*}$ , but it does not remove any item from it yet.

If bundle  $Z_{j^*}$  belongs to the augmenting path and appears in the edge  $(j^o, Z_{j^*})$  of  $M$ , the algorithm modifies  $M$  by removing edge  $(j^o, Z_{j^*})$  and adding edge  $(j_k, Z_{j^*})$  to it (Line 17). We will see shortly that this is a valid modification that does not violate the properties (a), (b), and (c) of matching  $M$ . The algorithm repeats the augmenting path process until the computed robust demand bundle  $Z_{j^*}$  does not belong to the augmenting path. In this case, the execution flow enters the else statement and the least valued item for agent  $j_k$  is removed from bundle  $Z_{j^*}$  (Line 19). Before completing the current round, the algorithm checks whether the value of agent  $j^*$  for bundle  $Z_{j^*}$  has dropped significantly below half (more than by a factor of  $2 + \delta_1$ ) of her value for the initial bundle  $X_{j^*}$  (Line 20). If this is the case, the algorithm computes a new allocation  $\widehat{\mathbf{X}}$  (Lines 21-25) which, as we will prove, has Nash welfare at least  $(1 + \delta)^{1/n} \text{NW}(\mathbf{X})$ . It then terminates, returning allocation  $\widehat{\mathbf{X}}$  as output. Otherwise, it indicates the end of the item removal process (Line 28), which will allow execution flow to leave the inner repeat-until loop.

**Input:** allocation  $\mathbf{X} = (X_1, \dots, X_n)$  of  $\mathcal{M}$  and  $\delta \in \mathbb{R}_+$ .

**Output:** allocation  $\mathbf{Y} = (Y_1, \dots, Y_n)$  of  $S \subseteq \mathcal{M}$  or allocation  $\widehat{\mathbf{X}} = (\widehat{X}_1, \dots, \widehat{X}_n)$  of  $\mathcal{M}$ .

- 1 Let  $\mathbf{Z} = (Z_1, \dots, Z_n) \leftarrow (X_1, \dots, X_n)$  be an ordered partition of  $\mathcal{M}$ ;
- 2 Let  $M_0 = \{(i, Z_i) : i \in [n]\}$ ;
- 3 Let  $\delta_1 = \frac{2\delta}{1-\delta}$ ;  $\triangleright \frac{2+2\delta_1}{2+\delta_1} = 1 + \delta$
- 4 **repeat**
- 5      $G \leftarrow \text{EFX-feasibility-graph}([n], \mathbf{Z})$ ;
- 6     Let  $T$  be the set of touched bundles in  $\mathbf{Z}$ ;
- 7     Compute a matching  $M$  in  $G$  such that
- 8         **(a)** All bundles in  $T$  are matched in  $M$ ,
- 9         **(b)** Under (a),  $|M \cap M_0|$  is maximized, and
- 10         **(c)** Under (a) and (b),  $|M|$  is maximized;
- 11     Let  $Z_{j_i}$  be an unmatched bundle in  $M$ ;
- 12     **repeat**
- 13         itemremoved  $\leftarrow$  false;
- 14         Let  $P = \{(j_1, Z_{j_2}), \dots, (j_{k-1}, Z_{j_k})\} \leftarrow \text{augmenting-path}(Z_{j_1}, M \cup M_0)$ ;
- 15          $Z_{j^*} \leftarrow \text{robust-demand}(j_k, \mathbf{Z})$ ;
- 16         **if**  $\exists (j^o, Z_{j^*}) \in P$  **then**
- 17              $M \leftarrow M \setminus \{(j^o, Z_{j^*})\} \cup \{(j_k, Z_{j^*})\}$ ;
- 18         **else**
- 19              $\mathbf{Z} \leftarrow (Z_1, \dots, Z_{j^*-1}, Z_{j^*} \setminus \{c\}, Z_{j^*+1}, \dots, Z_n)$ , where  $c \in \{\text{argmin}_{g \in Z_{j^*}} v_{j_k}(g)\}$ ;
- 20             **if**  $(2 + \delta_1)v_{j^*}(Z_{j^*}) < v_{j^*}(X_{j^*})$  **then**
- 21                  $\widehat{X}_{j_1} \leftarrow X_{j_1} \cup Z_{j_2}$ ;
- 22                  $\widehat{X}_{j_i} \leftarrow X_{j_i} \setminus Z_{j_i} \cup Z_{j_{i+1}}$  for  $i = 2, \dots, k-1$ ;
- 23                  $\widehat{X}_{j_k} \leftarrow X_{j_k} \setminus Z_{j_k} \cup Z_{j^*}$ ;
- 24                  $\widehat{X}_{j^*} \leftarrow X_{j^*} \setminus Z_{j^*}$ ;
- 25                  $\widehat{X}_i \leftarrow X_i$  for  $i \in [n] \setminus \{j_1, j_2, \dots, j_k, j^*\}$ .
- 26             **return**  $\widehat{\mathbf{X}}$ ;
- 27         **end**
- 28         itemremoved  $\leftarrow$  true;
- 29     **end**
- 30     **until** itemremoved;
- 31 **until**  $|M| = n$ ;
- 32 **return**  $\mathbf{Y} = (M_1, \dots, M_n)$ ;

**Algorithm 2:** Computes an EFX allocation of high Nash welfare or indicates that a significant improvement to the Nash welfare of the input allocation is possible.

For the analysis, we first claim that Lemma 1 carries over to the modified algorithm. All we need to show is that matching  $M$  keeps all its properties (a),(b), and (c) after every modification (in Line 17). As the modification affects neither the set of matched bundles, nor the size of  $M \cap M^o$ , it suffices to show that the new edge  $(j_k, Z_{j^*})$  does exist in the EFX feasibility graph. The proof is similar to the proof of Claim 4. Notice that the edge  $(j_k, Z_{j_k})$  is not EFX feasible, since otherwise it could replace edge  $(j_{k-1}, Z_{j_k})$  in  $M$  to increase  $|M \cap M^o|$ , still matching all touched bundles. Hence,

there exists  $j$  such that

$$v_{j_k}(Z_{j_k}) < \max_{g \in Z_j} v_{j_k}(Z_j \setminus \{g\}) \leq \max_{g \in Z_{j^*}} v_{j_k}(Z_{j^*} \setminus \{g\}) \leq v_{j_k}(Z_{j^*}),$$

where the second inequality follows since  $Z_{j^*}$  is the robust demand of agent  $j_k$ . Hence, the edge  $(j_k, Z_{j^*})$  indeed belongs to the EFX feasibility graph.

Now, if the algorithm terminates by returning allocation  $\mathbf{Y}$  in Line 32, this will be an EFX allocation satisfying  $(2 + \delta_1) \cdot v_i(Y_i) \geq v_i(X_i)$  for every agent  $i \in [n]$ . Furthermore, the argument in the proof of Claim 6 carries over, and one of the bundles, say  $X_{i^o}$ , will stay untouched until the end of the execution so that  $v_i(Y_{i^o}) \geq v_{i^o}(X_{i^o})$ . Consequently, the Nash welfare of allocation  $\mathbf{Y}$  is

$$(2 + \delta_1)^{1-1/n} \cdot \text{NW}(\mathbf{Y}) = \left( v_{i^o}(Y_{i^o}) \cdot \prod_{i \in [n] \setminus \{i^o\}} (2 + \delta_1) v_i(Y_i) \right)^{1/n} \geq \left( \prod_{i \in [n]} v_i(X_i) \right)^{1/n} = \text{NW}(\mathbf{X}).$$

This is summarized in the following statement.

**LEMMA 3.** *If the algorithm terminates and outputs the EFX allocation  $\mathbf{Y}$ , then  $(2 + \delta_1)^{1-1/n} \cdot \text{NW}(\mathbf{Y}) \geq \text{NW}(\mathbf{X})$ .*

If, in contrast, the execution of the algorithm enters Lines 21-25 and the algorithm terminates returning the allocation  $\widehat{\mathbf{X}}$ , we can show the following.

**LEMMA 4.** *If the algorithm terminates and outputs allocation  $\widehat{\mathbf{X}}$ , then  $\text{NW}(\widehat{\mathbf{X}}) \geq (1 + \delta)^{1/n} \text{NW}(\mathbf{X})$ .*

**PROOF.** When the execution enters lines 21-26, it holds that  $(2 + \delta_1) v_{j^*}(Z_{j^*}) < v_{j^*}(X_{j^*})$ , which implies that  $(2 + \delta_1) v_{j^*}(\widehat{X}_{j^*}) > (1 + \delta_1) v_{j^*}(X_{j^*})$ . As  $v_{j_i}(Z_{j_i}) \leq v_{j_i}(Z_{j_{i+1}})$  for  $i = 2, \dots, k-1$  and  $v_{j_k}(Z_{j_k}) \leq v_{j_k}(Z_{j^*})$ , we also have  $v_i(\widehat{X}_i) \geq v_i(X_i)$  for  $i \in [n] \setminus \{j_1, j^*\}$ . Furthermore,  $v_{j_1}(\widehat{X}_{j_1}) \geq 2v_{j_1}(X_{j_1})$ . Putting the above inequalities together, we have

$$\text{NW}(\widehat{\mathbf{X}}) = \left( \prod_{i \in [n]} v_i(\widehat{X}_i) \right)^{1/n} \geq \left( \frac{2(1 + \delta_1)}{2 + \delta_1} \right)^{1/n} \left( \prod_{i \in [n]} v_i(X_i) \right)^{1/n} = (1 + \delta)^{1/n} \text{NW}(\mathbf{X}).$$

The last equality is due to the relation of parameters  $\delta$  and  $\delta_1$  (Line 3).  $\square$

We keep running Algorithm 2, starting with the initial input allocation  $\mathbf{X}$ , and, every time it outputs an allocation  $\widehat{\mathbf{X}}$  with significantly higher Nash welfare than  $\mathbf{X}$ , we let  $\mathbf{X} \leftarrow \widehat{\mathbf{X}}$  and invoke it again. We stop when Algorithm 2 outputs an EFX allocation  $\mathbf{Y}$  for the first time. By Lemma 3, this is a  $(2 + \delta_1)^{1-1/n}$  approximation to the Nash welfare of the current input allocation  $\mathbf{X}$ . As the Nash welfare of the input allocation only improves with time, we get  $(2 + \delta_1)^{1-1/n} \cdot \rho \cdot \text{NW}(\mathbf{Y}) \geq \text{opt}(\mathcal{M})$ . Notice that no more than  $O\left(\frac{n\rho}{\delta}\right)$  executions of the algorithm will be required before Algorithm 2 outputs EFX allocation  $\mathbf{Y}$ . Indeed, by Lemma 4, we know that, after every  $n$  updates of  $\mathbf{X} \leftarrow \widehat{\mathbf{X}}$ , the Nash welfare of  $\mathbf{X}$  increases by a factor of at least  $(1 + \delta)$ , yielding an additive improvement to the Nash welfare of at least  $\delta \cdot \text{NW}(\mathbf{X}) \geq \frac{\delta}{\rho} \text{opt}(\mathcal{M})$ . Thus, if the Algorithm 2 does not output  $\mathbf{Y}$ ,  $\text{NW}(\mathbf{X})$  would become larger than  $\text{opt}(\mathcal{M})$  after  $O\left(\frac{n\rho}{\delta}\right)$  updates  $\mathbf{X} \leftarrow \widehat{\mathbf{X}}$ . Setting  $\delta = \frac{1}{2n+1}$ , and, consequently,  $\delta_1 = \frac{1}{n}$  we obtain a  $2\rho$ -approximation<sup>5</sup> after at most  $O(n^2)$  executions of Algorithm 2. The following statement summarizes the discussion of this section.

**THEOREM 3.** *On input a  $\rho$ -efficient allocation, Algorithm 2 returns a  $2\rho$ -efficient EFX allocation after at most  $O(n^2)$  repeated executions.*

<sup>5</sup>We get  $(2 + \delta_1)^{1-1/n} \cdot \rho$ -efficiency for the allocation  $\mathbf{Y}$  (by Lemma 3), where  $(2 + \frac{1}{n})^{1-1/n} < 2$ .

## 6 LARGE MARKETS

In this section, we give an improved guarantee for the Nash social welfare of the EFX allocation  $\mathbf{Y}$  produced by Algorithm 1, if the initial Nash social welfare maximizing allocation  $\mathbf{X}$  satisfies the large market condition with a parameter  $\varepsilon$ .

**THEOREM 4.** *If the input allocation  $\mathbf{X}$  satisfies the large market condition with a parameter  $\varepsilon$ , then Algorithm 1 outputs a  $(1 + 8\sqrt{\varepsilon})$ -efficient EFX allocation  $\mathbf{Y}$ .*

**PROOF.** In order to show the stated guarantee, we prove a stronger version of Lemma 2 where, instead of the guarantee  $v_i(X_i) \leq 2 \cdot v_i(Y_i)$ , we show that  $v_i(X_i) \leq (1 + 8 \cdot \sqrt{\varepsilon}) v_i(Y_i)$ , for each agent  $i \in [n]$ . We follow the approach in the proof of Lemma 2, with the only difference that we stop at the iteration  $t$  of Algorithm 1 when the condition

$$\forall i \in [n] \quad (1 + 8 \cdot \sqrt{\varepsilon}) v_i(Z_i(t)) \geq v_i(X_i) \quad (3)$$

is violated for the first time (instead of the condition  $2 \cdot v_i(Z_i(t)) \geq v_i(X_i)$  that we considered in the proof of Lemma 2).

As in Lemma 2, if condition (3) is violated for the bundle  $Z_{j^*}$ , we can find an augmenting path  $P$  defined in equation (2) with a set of vertices  $\{j_1, \dots, j_k\} \subset \mathcal{N}$  and  $\{Z_{j_1}, \dots, Z_{j_{k+1}}\} \subseteq \{Z_i\}_{i=1}^n$  from an unallocated bundle  $Z_{j_1}$  to the violating bundle  $Z_{j_{k+1}} = Z_{j^*}$ . To simplify notation, we assume without loss of generality that  $j_i = i$  for each  $i \in [k + 1]$ . Hence, we have

$$P \triangleq \{(1, Z_2), \dots, (k, Z_{k+1})\} \in M', \quad \{(1, Z_1), \dots, (k, Z_k)\} \in M^o.$$

We recall some key facts about the bundles in the augmenting path  $P$ . Since  $(i, Z_{i+1})$  is an edge in the EFX feasibility graph  $G$ ,  $v_i(Z_{i+1}) > v_i(Z_i)$  for all  $i \in [k]$ . Since condition (3) is violated for the bundle  $Z_{k+1}$ ,  $(1 + 8 \cdot \sqrt{\varepsilon}) v_{k+1}(Z_{k+1}) < v_{k+1}(X_{k+1})$ . Also, since bundle  $Z_1$  is unallocated, it must be untouched, i.e.,  $Z_1 = X_1$ .

Next, we consider how to reallocate items between each consecutive pair of bundles  $Z_i$  and  $Z_{i+1}$ . Unlike in the proof of Lemma 2, the reallocation of the entire set  $Z_{i+1}$  to each agent  $i$  does not necessarily increase the Nash social welfare. To obtain an improvement, we have to be more flexible and explore different reallocation schemes from agent  $i + 1$  to agent  $i$ . Among many possible schemes, we employ the greedy strategy. Specifically, we rename the items in  $Z_{i+1}$  so that  $Z_{i+1} = \{g_1, \dots, g_\ell\}$  with<sup>6</sup>

$$\frac{v_i(g_1)}{v_{i+1}(g_1)} \geq \frac{v_i(g_2)}{v_{i+1}(g_2)} \geq \dots \geq \frac{v_i(g_\ell)}{v_{i+1}(g_\ell)}.$$

In the greedy scheme, we consider to move the first  $r$  items, i.e., the set  $R(r) = R_{i+1}(r) \stackrel{\text{def}}{=} \{g_1, \dots, g_r\}$ , from agent  $i + 1$  to agent  $i$ . The following Claim establishes an important property of any such move: relative increase in agent  $i$ 's value is greater than the relative decrease in agent  $i + 1$ 's value.

**CLAIM 7.**  $\frac{v_i(R(r))}{v_i(Z_i)} \geq \frac{v_{i+1}(R(r))}{v_{i+1}(Z_{i+1})}$  for any  $r \in \{0, \dots, \ell\}$ .

**PROOF.** As  $v_i(Z_{i+1}) > v_i(Z_i)$ , we have  $\frac{v_i(R(r))}{v_i(Z_i)} \geq \frac{v_i(R(r))}{v_i(Z_{i+1})}$ . It remains to show that  $\frac{v_i(R(r))}{v_i(Z_i)} \geq \frac{v_{i+1}(R(r))}{v_{i+1}(Z_{i+1})}$  or, equivalently,  $v_i(R(r))v_{i+1}(Z_{i+1}) \geq v_i(Z_{i+1})v_{i+1}(R(r))$ . Let  $S = Z_{i+1} \setminus R(r)$ ; then we need to show  $v_i(R(r))(v_{i+1}(S) + v_{i+1}(R(r))) \geq (v_i(S) + v_i(R(r)))v_{i+1}(R(r))$  or, equivalently,  $x \stackrel{\text{def}}{=} v_i(R(r))v_{i+1}(S) - v_i(S)v_{i+1}(R(r)) \geq 0$ . By the definition of  $x$ ,  $R(r)$ , and the renaming of items in

<sup>6</sup>For simplicity of notation, in this section, we extensively use fractions that involve valuations of items or bundles in the numerator and the denominator. As we have allowed valuations of zero, we use the convention that the outcome of the comparison of two fractions with zeros in their denominators is the outcome of the comparison of their numerators.

$Z_{i+1}$ , we have that  $x = \sum_{j=1}^r \sum_{s=r+1}^\ell (v_i(g_j)v_{i+1}(g_s) - v_i(g_s)v_{i+1}(g_j))$ . Since  $\frac{v_i(g_j)}{v_{i+1}(g_j)} \geq \frac{v_i(g_s)}{v_{i+1}(g_s)}$  for any  $j \geq r \geq s$ , we have  $v_i(g_j)v_{i+1}(g_s) - v_i(g_s)v_{i+1}(g_j) \geq 0$ . Therefore,  $x \geq 0$ .  $\square$

Claim 7 gives us a powerful tool to balance the valuations of agent  $i$  and  $i+1$ . The major challenge, however, is to *simultaneously* adjust the valuations of the agents  $i \in [k+1]$  in a way that improves the Nash social welfare of the initial allocation  $\mathbf{X}$ . We consider arbitrary simultaneous reallocation of items between all pairs of agents  $i$  and  $i+1$ , where the first  $r_{i+1}$  items are moved from agent  $i+1$  to agent  $i$  according to the greedy scheme for all  $i \in [k]$ . We can focus only on the changes of the allocations for the first  $k+1$  agents, as the remaining bundles  $X_i$  remain unchanged for  $i \notin [k+1]$ . The new allocation  $\widehat{\mathbf{X}}$  for the first  $k+1$  agents is as follows.

$$\widehat{\mathbf{X}} : \begin{cases} \widehat{X}_1 & \stackrel{\text{def}}{=} X_1 \cup R_2(r_2) \\ \widehat{X}_i & \stackrel{\text{def}}{=} X_i \setminus R_i(r_i) \cup R_{i+1}(r_{i+1}), \quad i \in \{2, \dots, k\} \\ \widehat{X}_{k+1} & \stackrel{\text{def}}{=} X_{k+1} \setminus R_{k+1}(r_{k+1}) \end{cases}$$

We compare the Nash social welfare of the new allocation  $\widehat{\mathbf{X}}$  with that of the initial allocation  $\mathbf{X}$ .

$$\ln \left( \frac{\text{NW}(\widehat{\mathbf{X}})^n}{\text{NW}(\mathbf{X})^n} \right) = \ln \left( \prod_{i=1}^{k+1} \frac{v_i(\widehat{X}_i)}{v_i(X_i)} \right) = \sum_{i=1}^{k+1} \ln \left( \frac{v_i(\widehat{X}_i)}{v_i(Z_i)} \right) - \sum_{i=1}^{k+1} \ln \left( \frac{v_i(X_i)}{v_i(Z_i)} \right). \quad (4)$$

As every bundle  $Z_i$  is always a subset of the initial bundle  $X_i$ , we have  $v_i(X_i) = v_i(Z_i) + v_i(X_i \setminus Z_i)$  and  $\frac{v_i(X_i)}{v_i(Z_i)} = 1 + \frac{v_i(X_i \setminus Z_i)}{v_i(Z_i)}$ . Let  $a_i \stackrel{\text{def}}{=} \frac{v_i(X_i \setminus Z_i)}{v_i(Z_i)}$ , so  $\frac{v_i(X_i)}{v_i(Z_i)} = 1 + a_i$  for each  $i \in [k+1]$ . To simplify notation, let  $R_1(r_1), R_{k+2}(r_{k+2}) \stackrel{\text{def}}{=} \emptyset$ , so that  $\widehat{X}_i = X_i \setminus R_i(r_i) \cup R_{i+1}(r_{i+1})$  for each  $i \in [k+1]$  and  $\frac{v_i(\widehat{X}_i)}{v_i(Z_i)} = 1 + \frac{v_i(X_i \setminus Z_i)}{v_i(Z_i)} - \frac{v_i(R_i(r_i))}{v_i(Z_i)} + \frac{v_i(R_{i+1}(r_{i+1}))}{v_i(Z_i)}$ . Let  $x_i \stackrel{\text{def}}{=} \frac{v_i(R_i(r_i))}{v_i(Z_i)}$  for  $i \in [k+1]$ ; then we get

$$\frac{v_i(\widehat{X}_i)}{v_i(Z_i)} = 1 + a_i - x_i + \frac{v_i(R_{i+1}(r_{i+1}))}{v_i(Z_i)} \geq 1 + a_i - x_i + \frac{v_{i+1}(R_{i+1}(r_{i+1}))}{v_{i+1}(Z_{i+1})} = 1 + a_i - x_i + x_{i+1},$$

where the inequality follows from Claim 7. Therefore, we can continue equation (4):

$$(4) \geq \sum_{i=1}^{k+1} \ln [1 + a_i - x_i + x_{i+1}] - \sum_{i=1}^{k+1} \ln [1 + a_i], \quad (5)$$

where  $a_1 = 0$  and  $x_1 = x_{k+2} = 0$ . Moreover, we have  $8\sqrt{\varepsilon} \cdot v_i(Z_i) \geq v_i(X_i \setminus Z_i)$  by the definition of  $a_i$  for each  $i \in [k]$  and  $Z_{k+1}$  is the bundle that violates this inequality. Thus  $a_i \triangleq \frac{v_i(X_i \setminus Z_i)}{v_i(Z_i)} \leq 8\sqrt{\varepsilon}$  for every  $i \in [k]$  and  $a_{k+1} > 8\sqrt{\varepsilon}$ . Note that we control each  $x_i = \frac{v_i(R_i(r_i))}{v_i(Z_i)}$  for  $i \in \{2, \dots, k+1\}$  by choosing the corresponding number of items  $r_i$  to be reallocated to agent  $i-1$ . Let  $\Theta_i = \{\theta_{i,0} < \dots < \theta_{i,\ell_i}\}$  be the set of real numbers that  $x_i$  can take for each  $i \in \{2, \dots, k+1\}$ . Then for each such  $i$  we have

CLAIM 8. *The set  $\Theta_i = \{\theta_{i,0} < \dots < \theta_{i,\ell_i}\}$  has  $\theta_{i,0} = 0$ ,  $\theta_{i,\ell_i} = 1$  and  $|\theta_{i,j} - \theta_{i,j-1}| \leq 2\varepsilon$  for all  $j$ .*

PROOF. Setting  $r_i = 0$  and  $r_i = |Z_i|$  yields  $x_i = 0$  and  $x_i = 1$ , respectively. To bound the gap between consecutive  $\theta_{i,j-1}, \theta_{i,j}$  it suffices to bound  $\frac{v_i(g)}{v_i(Z_i)} \leq 2\varepsilon$  for any item  $g \in Z_i$ . We have  $v_i(Z_i) \geq \frac{1}{2}v_i(X_i)$  (by Lemma 2). Thus  $\frac{v_i(g)}{v_i(Z_i)} \leq 2\frac{v_i(g)}{v_i(X_i)} \leq 2\varepsilon$  by the large market assumption.  $\square$

Finally, we show that there are feasible  $x_i$ 's, so that the RHS of (5) is strictly positive.

LEMMA 5. For any sequence of numbers  $(a_i)_{i=1}^{k+1}$  such that  $a_1 = 0$ ,  $a_{k+1} \geq 8\sqrt{\varepsilon}$ , and  $a_i \in [0, 8\sqrt{\varepsilon}]$  there is a feasible solution  $(x_i \in \Theta_i)_{i=1}^{k+2}$  with  $x_1 = x_{k+2} = 0$  such that

$$\text{RHS (5)} = \sum_{i=1}^{k+1} \ln[1 + a_i - x_i + x_{i+1}] - \sum_{i=1}^{k+1} \ln[1 + a_i] > 0$$

PROOF. To prove the lemma, we use Karamata's inequality for the concave function  $\phi(x) = \ln(1 + x)$ . Namely, we find a sequence of feasible  $x_i$ 's such that the ordered sequence  $b_{(1)} \geq b_{(2)} \geq \dots \geq b_{(k+1)}$  with  $b_i = a_i - x_i + x_{i+1}$  for  $i \in [k+1]$  is (strictly) majorized by the ordered sequence  $a_{(1)} \geq a_{(2)} \geq \dots \geq a_{(k+1)}$ , i.e.,  $\sum_{i=1}^s b_i \leq \sum_{i=1}^s a_i$  for any  $1 \leq s \leq k+1$ , with equality for  $s = k+1$ . Then,  $\sum_{i=1}^{k+1} \phi(b_i) > \sum_{i=1}^{k+1} \phi(a_i)$  for any strictly concave function  $\phi$ .

First, we consider the case when there are consecutive  $a_{i-1}$  and  $a_i$  such that  $a_i - a_{i-1} > 2\varepsilon$ . In this case, we set  $x_j = 0$  for all  $j \neq i$  and  $x_i = \theta_{i,1}$ , i.e., the smallest non zero value in  $\Theta_i$ . Then almost all terms in both summations in the RHS of (5) are the same except for two terms  $\ln(1+a_{i-1}+x_i)+\ln(1+a_i-x_i)$  and  $\ln(1+a_{i-1})+\ln(1+a_i)$  in each summation. As  $x_i \leq 2\varepsilon$ ,  $b_{i-1} = a_{i-1}+x_i$  and  $b_i = a_i - x_i$  are closer to each other than  $a_{i-1}, a_i$  with the same sum  $b_{i-1} + b_i = a_{i-1} + a_i$ . Therefore, by Karamata's inequality  $\ln(1 + a_{i-1} + x_i) + \ln(1 + a_i - x_i) - \ln(1 + a_{i-1}) - \ln(1 + a_i) > 0$  and the proof of the lemma is complete.

So, in the following, we assume that  $a_i - a_{i-1} \leq 2\varepsilon$  for all  $i \in \{2, \dots, k+1\}$ . This means that  $a_{i+1} \leq 2\varepsilon \cdot i$  and  $a_{k+1-i} > 8\sqrt{\varepsilon} - 2\varepsilon \cdot i$  for any  $i \in [k+1]$ , as  $a_1 = 0$  and  $a_{k+1} > 8\sqrt{\varepsilon}$ . Let  $\ell \stackrel{\text{def}}{=} \left\lfloor \frac{2}{\sqrt{\varepsilon}} \right\rfloor$ , then  $a_1, \dots, a_{\ell+1} \leq 4\sqrt{\varepsilon} \leq a_{k-\ell+1}, \dots, a_{k+1}$ . First, we find an increasing sequence of  $(x_i)_{i=1}^{\ell+1}$ , such that  $x_1 = 0$ ,  $x_{\ell+1} = 1$  and each  $b_i$  becomes closer to the median  $4\sqrt{\varepsilon}$  than  $a_i$  (but still less than the median). To this end, we set  $x_1 = 0$ , and then, one after another, we set every following  $x_{i+1}$  equal to  $\text{argmax}_{\theta \in \Theta_{i+1}} \{\theta \mid a_i + \theta - x_i \leq 4\sqrt{\varepsilon}\}$ , where  $i \in [\ell]$ . Note that once the sequence of  $x_i$ 's reaches 1 it stays equal to 1, because maximal element  $\bar{\theta} \in \Theta_{i+1}$  is 1 and it satisfies the condition  $a_i + \bar{\theta} - 1 \leq 4\sqrt{\varepsilon}$  for all  $i \in [\ell]$ . Now, let us verify that sequence  $(x_i)_{i=1}^{\ell}$  reaches 1. Assume to the contrary that  $x_i < 1$  for  $i \leq \ell$ . That means by Claim 8 that each  $a_i - x_i + x_{i+1}$  must be at most  $2\varepsilon$  far from the median  $4\sqrt{\varepsilon}$ , as otherwise we would increase  $x_{i+1}$  and get closer to the median. Thus  $a_i - x_i + x_{i+1} > 4\sqrt{\varepsilon} - 2\varepsilon$  for all  $i \leq \ell + 1$ . Hence,

$$\begin{aligned} x_{\ell+1} &= (x_2 - x_1) + \dots + (x_{\ell+1} - x_\ell) > \sum_{i=1}^{\ell} (4\sqrt{\varepsilon} - 2\varepsilon - a_i) \geq \sum_{i=1}^{\ell} (4\sqrt{\varepsilon} - 2\varepsilon - 2\varepsilon(i-1)) \\ &= \sqrt{\varepsilon} \ell \cdot (4 - \sqrt{\varepsilon}(\ell+1)) \geq \sqrt{\varepsilon} \left( \frac{2}{\sqrt{\varepsilon}} - 1 \right) \cdot \left( 4 - \sqrt{\varepsilon} \left( \frac{2}{\sqrt{\varepsilon}} + 1 \right) \right) = (2 - \sqrt{\varepsilon})^2 \geq 1, \end{aligned}$$

where in the second inequality we used  $a_i \leq 2\varepsilon(i-1)$ , and in the third inequality we used  $\frac{2}{\sqrt{\varepsilon}} \geq \ell = \left\lfloor \frac{2}{\sqrt{\varepsilon}} \right\rfloor \geq \frac{2}{\sqrt{\varepsilon}} - 1$ . We get  $x_{\ell+1} > 1$ , which contradicts  $x_{\ell+1} \in \Theta_{\ell+1}$ . Now, it is also easy to see that sequence  $(x_i)_{i=1}^{\ell+1}$  is non-decreasing: either  $x_i$  has already reached 1 and stays constant, or  $a_i - x_i + x_{i+1} = b_i \geq \frac{2}{\sqrt{\varepsilon}} - 2\varepsilon \geq a_i$  for any  $i \leq \ell$ . Analogously, we find a decreasing sequence<sup>7</sup> of  $(x_i)_{i=k-\ell+1}^{k+2}$ , such that  $x_{k-\ell+1} = 1$ ,  $x_{k+2} = 0$  and each  $b_i$  becomes closer to the median  $4\sqrt{\varepsilon}$  than  $a_i$  (but still larger than  $4\sqrt{\varepsilon}$ ). We set the remaining  $x_i$  for  $\ell+1 < i < k-\ell+1$  to 1, which gives us  $b_i = a_i$  for  $\ell+1 < i < k-\ell+1$  (recall that  $x_{\ell+1} = x_{k-\ell+1} = 1$ ). Now, it is easy to see that  $\{b_i\}_{i=1}^{k+1}$  is majorized by  $\{a_i\}_{i=1}^{k+1}$  which concludes the proof of the lemma by Karamata's inequality.  $\square$

<sup>7</sup>The argument mirrors the one for  $i \in [\ell+1]$ , i.e., we set  $x_{k+2} = 0$  first and then let each previous  $x_{i-1} = \text{argmax}_{\theta \in \Theta_{i-1}} \{\theta \mid a_{i-1} - \theta + x_i \geq 4\sqrt{\varepsilon}\}$

Lemma 5 implies that Algorithm 1 must terminate before any quantity  $(1 + 8\sqrt{\varepsilon})v_i(Z_i)$  becomes smaller than  $v_i(X_i)$ , since otherwise there is an allocation  $\widehat{X}$  with a higher Nash social welfare than the optimal one of  $X$ . Moreover, for any edge  $(i, Z_j)$  in the EFX-feasibility-graph  $v_i(Z_j) \geq v_i(Z_i)$ . Hence, Algorithm 1 for large markets outputs an EFX allocation  $Y$ , with  $(1 + 8\sqrt{\varepsilon})NW(Y) \geq (1 + 8\sqrt{\varepsilon})NW(Z) \geq NW(X) = \text{opt}(M)$ .  $\square$

## 7 DISCUSSION AND OPEN PROBLEMS

We believe that our techniques could be used to show interesting interplays of more fairness notions with Nash welfare. For example, starting from the allocation computed by Algorithm 2, one could use the local-search algorithm of Lipton et al. [29] to reallocate the removed items and get an EF1 allocation of *all* items. To the best of our knowledge, this is the first polynomial-time algorithm for computing a complete EF1 allocation that approximates maximum Nash welfare within a constant.

Still, the problem of whether EFX allocations of *all* items exist is widely open. If EFX allocations do not always exist, our proposed solution alleviates the existence issue while providing high efficiency guarantees. But we suspect that there is a monotonicity property that, if true, would not only show that EFX allocations always exist, but also that they are nearly-optimal in terms of Nash welfare. In particular, we suspect that adding an item to an allocation problem (that provably has an EFX allocation) yields another problem that also has an EFX allocation with at least as high Nash welfare as the initial one. Then, our Theorem 2 would imply not only the existence of EFX allocations for all items, but also that the best among them is  $2^{1-1/n}$ -efficient.

## ACKNOWLEDGMENTS

Part of this work was done while authors IC and XH were visiting the Institute for Theoretical Computer Science at Shanghai University of Finance and Economics.

## REFERENCES

- [1] Georgios Amanatidis, Georgios Birmpas, and Vangelis Markakis. 2018. Comparing approximate relaxations of envy-freeness. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 42–48.
- [2] Georgios Amanatidis, Evangelos Markakis, Afshin Nikzad, and Amin Saberi. 2017. Approximation algorithms for computing maximin share allocations. *ACM Transactions on Algorithms* 13, 4 (2017), 52:1–52:28.
- [3] Nima Anari, Gagan Goel, and Afshin Nikzad. 2014. Mechanism design for crowdsourcing: An optimal  $1-1/e$  competitive budget-feasible mechanism for large markets. In *Proceedings of the 55th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*. 266–275.
- [4] Nima Anari, Tung Mai, Shayan Oveis Gharan, and Vijay V Vazirani. 2018. Nash social welfare for indivisible items under separable, piecewise-linear concave utilities. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2274–2290.
- [5] Nima Anari, Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. 2017. Nash social welfare, matrix permanent, and stable polynomials. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS)*. 36:1–12.
- [6] Haris Aziz, Sylvain Bouveret, Ioannis Caragiannis, Ira Giagkousi, and Jérôme Lang. 2018. Knowledge, fairness, and social constraints. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 4638–4645.
- [7] Haris Aziz and Simon Mackenzie. 2016. A discrete and bounded envy-free cake cutting protocol for any number of Agents. In *Proceedings of the 57th Annual Symposium on Foundations of Computer Science (FOCS)*. 416–427.
- [8] Eric Balkanski and Yaron Singer. 2015. Mechanisms for fair attribution. In *Proceedings of the 16th ACM Conference on Economics and Computation (EC)*. 529–546.
- [9] Siddharth Barman and Sanath Kumar Krishna Murthy. 2017. Approximation algorithms for maximin fair division. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*. 647–664.
- [10] Siddharth Barman, Sanath Kumar Krishnamurthy, and Rohit Vaish. 2018. Finding fair and efficient allocations. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*. 557–574.
- [11] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. 2011. The price of fairness. *Operations Research* 59, 1 (2011), 17–31.
- [12] Sylvain Bouveret and Michel Lemaître. 2016. Characterizing conflicts in fair division of indivisible goods using a scale of criteria. *Autonomous Agents and Multi-Agent Systems* 30, 2 (2016), 259–290.

- [13] Eric Budish. 2011. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103.
- [14] Ioannis Caragiannis, Christos Kaklamani, Panagiotis Kanellopoulos, and Maria Kyropoulou. 2012. The efficiency of fair division. *Theory of Computing Systems* 50, 4 (2012), 589–610.
- [15] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D Procaccia, Nisarg Shah, and Junxing Wang. 2016. The unreasonable fairness of maximum Nash welfare. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*. 305–322.
- [16] Bhaskar Ray Chaudhury, Yun Kuen Cheung, Jugal Garg, Naveen Garg, Martin Hoefer, and Kurt Mehlhorn. 2018. On Fair Division for Indivisible Items. In *38th LARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- [17] Yann Chevaleyre, Ulle Endriss, Sylvia Estivie, and Nicolas Maudet. 2007. Reaching envy-free states in distributed negotiation settings. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*. 1239–1244.
- [18] Richard Cole, Nikhil R Devanur, Vasilis Gkatzelis, Kamal Jain, Tung Mai, Vijay V Vazirani, and Sadra Yazdanbod. 2017. Convex program duality, Fisher markets, and Nash social welfare. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*. 459–460.
- [19] Richard Cole and Vasilis Gkatzelis. 2015. Approximating the Nash social welfare with indivisible items. In *Proceedings of the 47th Annual ACM on Symposium on Theory of Computing (STOC)*. 371–380.
- [20] Richard Cole and Yixin Tao. 2016. Large market games with near optimal efficiency. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*. 791–808.
- [21] E. Eisenberg and D. Gale. 1959. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics* 30, 1 (1959), 165–168.
- [22] Michal Feldman, Nick Gravin, and Brendan Lucier. 2016. Combinatorial Walrasian equilibrium. *SIAM J. Comput.* 45, 1 (2016), 29–48.
- [23] Michal Feldman, Nicole Immorlica, Brendan Lucier, Tim Roughgarden, and Vasilis Syrgkanis. 2016. The price of anarchy in large games. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*. 963–976.
- [24] Jugal Garg, Martin Hoefer, and Kurt Mehlhorn. 2018. Approximating the Nash social welfare with budget-additive valuations. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2326–2340.
- [25] Mohammad Ghodsi, MohammadTaghi HajiAghayi, Masoud Seddighin, Saeed Seddighin, and Hadi Yami. 2018. Fair allocation of indivisible goods: Improvements and generalizations. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*. 539–556.
- [26] Jonathan R. Goldman and Ariel D. Procaccia. 2014. Spliddit: unleashing fair division algorithms. *SIGecom Exchanges* 13, 2 (2014), 41–46.
- [27] David Kurokawa, Ariel D. Procaccia, and Junxing Wang. 2018. Fair enough: Guaranteeing approximate maximin shares. *J. ACM* 65(2) (2018), 8:1–27.
- [28] Euiwoong Lee. 2017. APX-hardness of maximizing Nash social welfare with indivisible items. *Inform. Process. Lett.* 122 (2017), 17–20.
- [29] Richard J Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. 2004. On approximately fair allocations of indivisible goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce (EC)*. 125–131.
- [30] H. Moulin. 2003. *Fair Division and Collective Welfare*. MIT Press.
- [31] John F Nash Jr. 1950. The bargaining problem. *Econometrica: Journal of the Econometric Society* (1950), 155–162.
- [32] Nhan-Tam Nguyen, Trung Thanh Nguyen, Magnus Roos, and Jörg Rothe. 2014. Computational complexity and approximability of social welfare optimization in multiagent resource allocation. *Autonomous Agents and Multi-Agent Systems* 28, 2 (2014), 256–289.
- [33] Trung Thanh Nguyen and Jörg Rothe. 2014. Minimizing envy and maximizing average Nash social welfare in the allocation of indivisible goods. *Discrete Applied Mathematics* 179 (2014), 54–68.
- [34] Benjamin Plaut and Tim Roughgarden. 2018. Almost envy-freeness with general valuations. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2584–2603.
- [35] Hugo Steinhaus. 1948. The problem of fair division. *Econometrica* 16, 1 (1948), 101–104.