Algorithms and Complexity Group
Institute for Computer Science
Goethe University Frankfurt

Summer Semester 2023

Seminar Approximation Algorithms

ANSWuSVþ(U)M

Zeno Adrian Weil

24th May 2023

Examiner: Prof Dr Martin Hoefer
Supervisor: Dr Giovanna Varricchio

Abstract

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Todo list

1 Introduction

- problem introduction, motivation, applications
- formal problem definition (incl. why geometric mean?)
- short literature review: What is known, what not? New findings?
- content & structure of paper

Definition 1. Let $\mathcal{G} := \{1, \dots, m\}$ be a set of indivisible *items* and $\mathcal{A} := \{1, \dots, n\}$ be a set of *agents*. An *allocation* is a tuple $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathcal{P}(G)^n$ of *bundles* $\boldsymbol{x}_i$ such that each item is element of exactly one bundle, that is $\bigcup_{i \in \mathcal{A}} \boldsymbol{x}_i = \mathcal{G}$ and $\boldsymbol{x}_i \cap \boldsymbol{x}_{i'} = \emptyset$ for all $i \neq i'$. An item $j \in \mathcal{G}$ is *assigned* to agent $i \in \mathcal{A}$ if $j \in \boldsymbol{x}_i$ holds.

or rather 'allocated'?

⋮

Definition 2. Given a set $\mathcal{G}$ of items and a set $\mathcal{A}$ of agents with *valuations* $v_i : \mathcal{P}(\mathcal{G}) \to \mathbb{R}$ and *agent weights* η_i for all agents $i \in \mathcal{A}$, the *Nash Social Welfare problem* (NSW) is to find an allocation maximising the weighted geometric mean of valuations, that is

i: less strict def of valuations; restriction for our case later on

$$\operatorname*{arg\,max}_{\boldsymbol{x} \in \Pi_n(\mathcal{G})} \left\{ \left(\prod_{i \in \mathcal{A}} v_i(\boldsymbol{x}_i)^{\eta_i} \right)^{1/\sum_{i \in \mathcal{A}} \eta_i} \right\}$$

where $\Pi_n(\mathcal{G})$ is the set of all possible allocations of the items in $\mathcal{G}$ amongst n agents. The problem is called *symmetric* if all agent weights η_i are equal, and *asymmetric* otherwise.

agents without items assigned have valuation zero → prevent

⋮

In a slight abuse of notation, we omit curly braces delimiting a set in the arguments of a valuation function, so for example we write $v(j_1, j_2, \dots)$ to denote $v(\{j_1, j_2, \dots\})$.

⋮

definition of approximation factor [def environment or in-text?]

⋮

Garg, Kulkarni and Kulkarni [1] consider five different types of non-negative monotonically non-decreasing valuation functions of which we are going to consider only the following two due to space constraints:

Additive The valuation $v_i(\mathcal{S})$ of an agent i for a set $\mathcal{S} \subset \mathcal{G}$ of items j is the sum of individual valuations $v_i(j)$, that is $v_i(\mathcal{S}) = \sum_{j \in \mathcal{S}} v_i(j)$.

Submodular Let $v_i(\mathcal{S}_1 \mid \mathcal{S}_2) := v_i(\mathcal{S}_1 \cup \mathcal{S}_2) - v_i(\mathcal{S}_2)$ denote the marginal valuation of agent i for a set $\mathcal{S}_1 \subset \mathcal{G}$ of items over a *disjoint* set $\mathcal{S}_2 \subset \mathcal{G}$. This valuation functions satisfies the submodularity constraint $v_i(j \mid \mathcal{S}_1 \cup \mathcal{S}_2) \leq v_i(j \mid \mathcal{S}_1)$ for all agents $i \in \mathcal{A}$, items $j \in \mathcal{G}$ and sets $\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{G}$ of items.

or rather utility?

We use *additive NSW* and *submodular NSW* as shorthands for the Nash social welfare problems with additive and submodular valuation functions, respectively.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

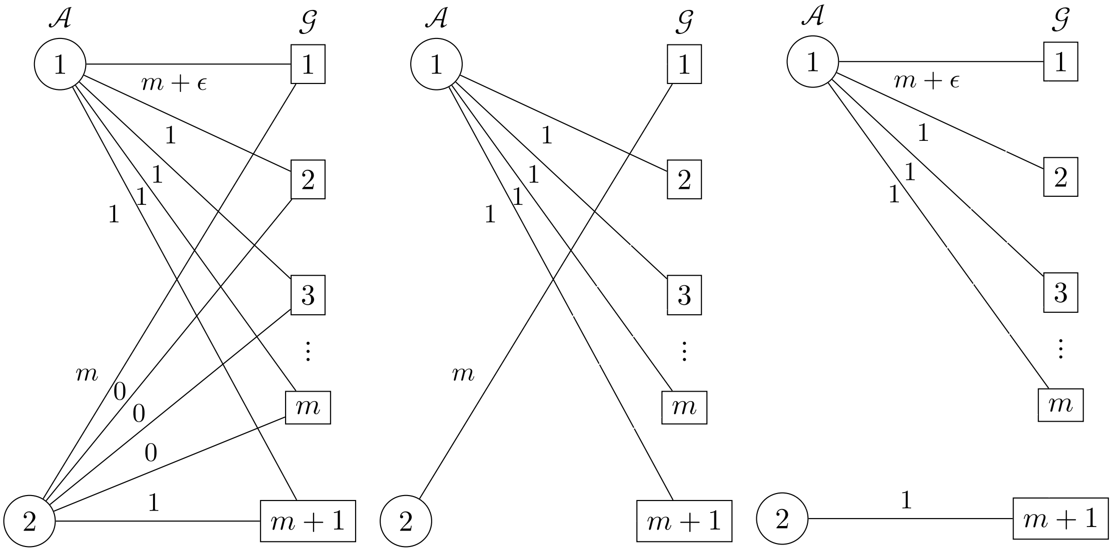
Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

Etiam ac leo a risus tristique nonummy. Donec dignissim tincidunt nulla. Vestibulum rhoncus molestie odio. Sed lobortis, justo et pretium lobortis, mauris turpis condimentum augue, nec ultricies nibh arcu pretium enim. Nunc purus neque, placerat id, imperdiet sed, pellentesque nec, nisl. Vestibulum imperdiet neque non sem accumsan laoreet. In hac habitasse platea dictumst. Etiam condimentum facilisis libero. Suspendisse in elit quis nisl aliquam dapibus. Pellentesque auctor sapien. Sed egestas sapien nec lectus. Pellentesque vel dui vel neque bibendum viverra. Aliquam porttitor nisl nec pede. Proin mattis libero vel turpis. Donec rutrum mauris et libero. Proin euismod porta felis. Nam lobortis, metus quis elementum commodo, nunc lectus elementum mauris, eget vulputate ligula tellus eu neque. Vivamus eu dolor.

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.



(a) problem instance **(b)** OPT $= \sqrt{m \cdot m}$ **(c)** ALG $\leq \sqrt{(2m+\epsilon-1)\cdot 1}$

i: might delete if document in two columns → unused edges in grey

Figure 1: Agent 1 values item 1 at $m+\epsilon$ and all other items at 1. Agent 2 values item 1 at m, item $m+1$ at 1 and all other items at 0. In an optimal allocation, item 1 would be assigned to agent 2 and all other items to agent 1, resulting in a NSW of $\sqrt{m \cdot m} = m$. A repeated maximum matching algorithm would greedily assign item 1 to agent 1 and item $m+1$ to agent 2 in the first round. Even if all remaining items were going to be assigned to agent 1, the NSW will never surpass $\sqrt{(2m+\epsilon-1)\cdot 1} < \sqrt{2m}$. The approximation factor $\alpha \approx \sqrt{m/2}$ therefore depends on the number of items.

2 SMatch

In the case of an equal number of agents and items, i. e., $|\mathcal{A}| = n = m = |\mathcal{G}|$, the *additive* NSW can be solved exactly by finding a maximum matching on a bipartite graph with the sets of agents and of items as its parts; as weight of the edge between agent i and item j, use $\eta_i \log v_i(j)$, that is the weighted valuation of item j by agent i in the *logarithmic* Nash social welfare. Should there be more items than agents, then it would be obvious to just repeatedly find a maximum matching and assign the items accordingly until all items are assigned. The flaw of this idea is that such a greedy algorithm only considers the valuations of items in the current matching and perhaps the valuations of items already assigned. As the example in fig. 1 demonstrates, this leads to an algorithm with an approximation factor dependent on the number m of items. The geometric mean of the NSW favours allocations with similarly valued bundles, wherefore it may be beneficial for an agent to leave items to other agents if those agents cannot expect many more valuable items in the future.

What is the factor?

The algorithm SMatch, described in algorithm 1, eliminates the flaw by first gaining foresight of the valuations of items assigned after the first matching, achieving an approximation factor of $2n$ (cf. theorem 1 later on). For a fixed agent i, order the items in descending order of the valuations by agent i and denote the j-th most liked item by $\mathcal{G}_i^j$. To obtain a well-defined order, items of equal rank are further ordered numerically. SMatch, too, does repeatedly match items. During the first matching, however, the edge weights are defined as $\eta_i \log(v_i(j) + u_i/n)$ for an edge between agent i and item j. The addend u_i serves as estimation of the valuation of items assigned after the first matching

Algorithm 1: SMatch for the Asymmetric Additive NSW problem

Input : set $\mathcal{A} = \{1, \dots, n\}$ of agents with weights $\eta_i \forall i \in \mathcal{A}$, set $\mathcal{G} = \{1, \dots, m\}$ of indivisible items, additive valuations $v_i \colon \mathcal{P}(\mathcal{G}) \to \mathbb{R}_0^+$ where $v_i(\mathcal{S})$ is the valuation of agent $i \in \mathcal{A}$ for each set $\mathcal{S} \subset \mathcal{G}$ of items

Output: $\frac{1}{2n}$-approximation $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ of an optimal allocation

$\boldsymbol{x}_i \leftarrow \emptyset \quad \forall i \in \mathcal{A}$

$u_i \leftarrow v_i(\mathcal{G}_i^{2n+1}, \dots, \mathcal{G}_i^m) \quad \forall i \in \mathcal{A}$ ▷ *est. valuations after the 1st matching*

$\mathcal{W} \leftarrow \{\, \eta_i \cdot \log(v_i(j) + u_i/n) \mid i \in \mathcal{A}, j \in \mathcal{G} \,\}$ ▷ *edge weights*

$G \leftarrow (\mathcal{A}, \mathcal{G}, \mathcal{W})$ ▷ *bipartite graph*

$\mathcal{M} \leftarrow \text{max\_weight\_matching}(G)$

$\boldsymbol{x}_i \leftarrow \{\, j \mid (i, j) \in \mathcal{M} \,\} \quad \forall i \in \mathcal{A}$ ▷ *assign according to matching*

$\mathcal{G}^{\text{rem}} \leftarrow \mathcal{G} \setminus \{\, j \mid (i, j) \in \mathcal{M} \,\}$ ▷ *remove assigned items*

while $\mathcal{G}^{\text{rem}} \neq \emptyset$ **do**

  $\mathcal{W} \leftarrow \{\, \eta_i \cdot \log(v_i(j) + v_i(\boldsymbol{x}_i)) \mid i \in \mathcal{A}, j \in \mathcal{G}^{\text{rem}} \,\}$

  $G \leftarrow (\mathcal{A}, \mathcal{G}^{\text{rem}}, \mathcal{W})$

  $\mathcal{M} \leftarrow \text{max\_weight\_matching}(G)$

  $\boldsymbol{x}_i \leftarrow \boldsymbol{x}_i \cup \{\, j \mid (i, j) \in \mathcal{M} \,\} \quad \forall i \in \mathcal{A}$

  $\mathcal{G}^{\text{rem}} \leftarrow \mathcal{G}^{\text{rem}} \setminus \{\, j \mid (i, j) \in \mathcal{M} \,\}$

end while

return $\boldsymbol{x}$

and is defined as

$$u_i := \min_{\substack{\mathcal{S} \subset \mathcal{G} \\ |\mathcal{S}| \leq 2n}} \{v_i(\mathcal{G} \setminus \mathcal{S})\} = v_i(\mathcal{G}_i^{2n+1}, \dots, \mathcal{G}_i^m). \tag{1}$$

The set $\mathcal{S}$ has less than $2n$ elements only if there are less than $2n$ items in total. From the second matching onwards, the edge weights are defined as $\eta_i \log(v_i(j) + v_i(\boldsymbol{x}_i))$, where $\boldsymbol{x}_i$ is the continuously updated bundle of agent i. The addend $v_i(\boldsymbol{x}_i)$ could lead to better allocations in applications, but does not improve the approximation factor asymptotically.

For convenience, we order the items in the final bundle $\boldsymbol{x}_i = \{h_i^1, \dots, h_i^{\tau_i}\}$ of agent i by the order in which they were assigned, so that item h_i^t is assigned according to the t-th matching. To calculate the approximation factor of SMatch, we first need to establish a lower bound on the valuation of single items.

Lemma 1. *For each agent $i \in \mathcal{A}$ and her final bundle $\boldsymbol{x}_i = \{h_i^1, \dots, h_i^{\tau_i}\}$, the item h_i^t is worth at least as much as the overall tn-th most highly valued item $\mathcal{G}_i^{tn}$, i. e. $v_i(h_i^t) \geq v_i(\mathcal{G}_i^{tn})$, for all $t = 1, \dots, \tau_i$.*

Proof. At the start of the t-th round, no more than $(t-1)n$ items out of the tn most highly valued items $\mathcal{G}_i^1, \dots, \mathcal{G}_i^{tn}$ have been assigned in previous rounds since at most n items are assigned in each iteration. During the t-th round, at most $n-1$ more of those highly valued items could be assigned to all other agents $i' \neq i$, leaving at least one item in $\mathcal{G}_i^1, \dots, \mathcal{G}_i^{tn}$ unassigned. Since $v_i(\mathcal{G}_i^k) \geq v_i(\mathcal{G}_i^{tn})$ for all $k \leq tn$ by definition of $\mathcal{G}_i^{\cdot}$, the lemma follows. □

We can now establish $u_i/n = v_i(\mathcal{G}_i^{2n+1}, \dots, \mathcal{G}_i^m)/n$ as lower bound on the valuations of items assigned after the first matching.

Lemma 2. *For each agent $i \in \mathcal{A}$ and her final bundle $\boldsymbol{x}_i = \{h_i^1, \dots, h_i^{\tau_i}\}$, it holds $v_i(h_i^2, \dots, h_i^{\tau_i}) \geq u_i/n$, where n is the number of items and $u_i = v_i(\mathcal{G}_i^{2n+1}, \dots, \mathcal{G}_i^m)$.*

Proof. By lemma 1 and definition of $\mathcal{G}_i^{\cdot}$, every item h_i^t is worth at least as much as each item $\mathcal{G}_i^{tn+k}$ with $k \in \{0, \dots, n-1\}$ and, consequently, its valuation $v_i(h_i^t)$ is at least as high as the mean valuation $\frac{1}{n} v_i(\mathcal{G}_i^{tn}, \dots, \mathcal{G}_i^{tn+n-1})$. Further, it holds $\tau_i n + n \geq m$ since each agent receives items for at least $\lfloor \frac{m}{n} \rfloor \geq \frac{m}{n} - 1$ rounds. Together, this yields

$$v_i(h_i^2, \dots, h_i^{\tau_i}) = \sum_{t=2}^{\tau_i} v_i(h_i^t) \geq \sum_{t=2}^{\tau_i} \frac{1}{n} v_i(\mathcal{G}_i^{tn}, \dots, \mathcal{G}_i^{tn+n-1}) \tag{2}$$

$$\geq \frac{1}{n} v_i(\mathcal{G}_i^{2n}, \dots, \mathcal{G}_i^{m-1}) \geq \frac{1}{n} v_i(\mathcal{G}_i^{2n+1}, \dots, \mathcal{G}_i^{m}) = \frac{u_i}{n} \tag{3}$$

with the last inequality stemming from $v_i(\mathcal{G}_i^{2n}) \geq v_i(\mathcal{G}_i^m)$. □

Remark 1. In lemma 2, we assumed non-zero valuations for all items, hence the bundle lengths of $\tau_i \geq \lfloor \frac{m}{n} \rfloor$. Of course in an actual program, one would not assign items to agents who value them at zero. Inasmuch as additional zeros in eq. (3) do not change the sum, lemma 2 still holds nevertheless.

This allows us to calculate an approximation factor for SMatch by comparing its output with an optimal allocation $\boldsymbol{x}^*$.

Theorem 1. *SMatch has an approximation factor of $2n$.*

Proof. Lemma 2 can be plugged into the logarithmic NSW:

$$\log \mathrm{NSW}(\boldsymbol{x}) = \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log v_i(h_i^1, \dots, h_i^{\tau_i}) \tag{4}$$

$$= \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log(v_i(h_i^1) + v_i(h_i^2, \dots, h_i^{\tau_i})) \tag{5}$$

$$\geq \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log(v_i(h_i^1) + u_i/n) \tag{6}$$

Notice that the first matching of SMatch maximises the sum in eq. (6). Thus, assigning all agents i their respective most highly valued item g_i^1 in an optimal bundle $\boldsymbol{x}_i^* = \{g_i^1, \dots, g_i^{\tau_i^*}\}$ yields the even lower bound

$$\log \mathrm{NSW}(\boldsymbol{x}) \geq \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log(v_i(g_i^1) + u_i/n). \tag{7}$$

Recall the definition of u_i from eq. (1). Consider a slightly modified variant:

$$u_i = \min_{\substack{\mathcal{S} \subset \mathcal{G} \\ |\mathcal{S}| \leq 2n}} \{v_i(\mathcal{G} \setminus \mathcal{S})\} \quad \text{or, alternatively,} \quad u_i = v_i(\mathcal{G} \setminus \mathcal{S}_i) \text{ with } \mathcal{S}_i := \arg\min_{\substack{\mathcal{S} \subset \mathcal{G} \\ |\mathcal{S}| \leq 2n}} \{v_i(\mathcal{G} \setminus \mathcal{S})\} \tag{8}$$

Moreover, consider the set $\mathcal{S}_i^*$ of the (at most) $2n$ most highly valued items in the optimal bundle $\boldsymbol{x}_i^*$, i. e.

$$\mathcal{S}_i^* := \arg\min_{\substack{\mathcal{S} \subset \mathcal{G} \\ |\mathcal{S}| \leq 2n}} \{v_i(\boldsymbol{x}_i^* \setminus \mathcal{S})\}. \tag{9}$$

We get the lower bound $v_i(g_i^1) \geq \frac{1}{2n} v_i(\mathcal{S}^*)$ from a similar argument as in the proof of lemma 2. Further, it holds $u_i = v_i(\mathcal{G} \setminus \mathcal{S}_i) \geq v_i(\boldsymbol{x}_i^* \setminus \mathcal{S}_i) \geq v_i(\boldsymbol{x}_i^* \setminus \mathcal{S}_i^*)$. We can substitute these two inequalities into eq. (7) and prove the theorem thereby:

$$\log \mathrm{NSW}(\boldsymbol{x}) \geq \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log\left(\frac{v_i(\mathcal{S}_i^*)}{2n} + \frac{v_i(\boldsymbol{x}_i^* \setminus \mathcal{S}_i^*)}{n}\right) \tag{10}$$

$$\geq \frac{1}{\sum_{i=1}^n \eta_i} \cdot \sum_{i=1}^n \eta_i \log\left(\frac{v_i(\boldsymbol{x}_i^*)}{2n}\right) = \log\left(\frac{\mathrm{NSW}(\boldsymbol{x}^*)}{2n}\right) \tag{11}$$

□

3 RepReMatch

- SMatch does not work for general submodular valuations since we need to detect the set of lowest valuation. This is not possible independent of m [SF11].

what exactly

First, we bound the valuation of an agent i for her optimal items which still may be assigned to her in phase II.

i: Would 'to bound' (strongly) imply giving an upper bound as well?

Lemma 3. *During phase II, each agent $i \in \mathcal{A}$ gets assigned a bundle $\boldsymbol{x}_i^{\mathrm{II}} = \{h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}}\}$. For all rounds $r = 2, \dots, \tau_i^{\mathrm{II}}$, the following inequality about her set $\bar{\boldsymbol{x}}_{i,r}^*$ of optimal and attainable items holds true:*

$$v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1}) \geq u_i^* - \ell_{i,2} v_i(h_i^1) - \sum_{r'=2}^{r-1} \ell_{i,r'} \cdot v_i(h_i^{r'} \mid h_i^1, \dots, h_i^{r'-1}) - v_i(h_i^1, \dots, h_i^{r-1}),$$

where $\ell_{i,l}$ is the size of her set $\mathcal{L}_{i,l} \subset \bar{\boldsymbol{x}}_{i,l-1}^*$ of optimal items which were attainable in round $l-1$ and were assigned to other agents in round l, and $u_i^* = v_i(\bar{\boldsymbol{x}}_{i,1}^*)$ is her valuation of attainable and optimal items during the first round.

Proof. We prove the lemma by induction on the number r of rounds. In the beginning of the base case $r = 2$, agent i has already been assigned item h_i^1 but not the items in set $\mathcal{L}_{i,1}$. For each of the remaining optimal and attainable items j in round 1, the marginal valuation $v_i(j \mid \emptyset)$ over the empty set is at most $v_i(h_i^1 \mid \emptyset)$, as otherwise item h_i^1 would not have been assigned first. Furthermore, the marginal valuation $v_i(j \mid h_i^1)$ over $\{h_i^1\}$ is upper-bounded by $v_i(h_i^1 \mid \emptyset)$ due to the submodularity of valuations. During the round, a further $\ell_{i,2}$ out of these items are assigned to other agents, and item h_i^2 is assigned to agent i. We can bound the marginal valuation of the remaining optimal and attainable items in round 2 in the following way:

Case 1 — $h_i^1 \in \bar{\boldsymbol{x}}_{i,1}^*$: It holds $v_i(\bar{\boldsymbol{x}}_{i,2}^* \mid h_i^1) = v_i(\bar{\boldsymbol{x}}_{i,2}^* \cup \{h_i^1\}) - v_i(h_i^1) = v_i(\bar{\boldsymbol{x}}_{i,1}^* \setminus \mathcal{L}_{i,2}) - v_i(h_i^1)$.

Case 2 — $h_i^1 \notin \bar{\boldsymbol{x}}_{i,1}^*$: Due to the monotonicity of valuations, it holds $v_i(\bar{\boldsymbol{x}}_{i,2}^* \cup \{h_i^1\}) \geq v_i(\bar{\boldsymbol{x}}_{i,2}^*)$ and, therefore, $v_i(\bar{\boldsymbol{x}}_{i,2}^* \mid h_i^1) \geq v_i(\bar{\boldsymbol{x}}_{i,2}^*) - v_i(h_i^1) = v_i(\bar{\boldsymbol{x}}_{i,1}^* \setminus \mathcal{L}_{i,2}) - v_i(h_i^1)$.

In both cases, the base case is proved because

$$\begin{aligned} v_i(\bar{\boldsymbol{x}}_{i,2}^* \mid h_i^1) &\geq v_i(\bar{\boldsymbol{x}}_{i,1}^* \setminus \mathcal{L}_{i,2}) - v_i(h_i^1) && (12)\\ &\geq v_i(\bar{\boldsymbol{x}}_{i,2}^*) - v_i(\mathcal{L}_{i,2}) - v_i(h_i^1) && (13)\\ &\geq u_i^* - \ell_{i,2} v_i(h_i^1) - v_i(h_i^1), && (14) \end{aligned}$$

where the second inequality can be shown inductively with the definition of submodularity, and the third inequality is due all $\ell_{i,2}$ items j in $\mathcal{L}_{i,2}$ being attainable but not assigned in round 1, implying $v_i(j) \leq v_i(h_i^1)$.

Do that or, alternatively, find a paper show-

As induction hypothesis, we assume that the lemma is true for all rounds up to some r. For the induction step $r \to r+1$, we differentiate the same two cases again:

Case 1 — $h_i^r \in \bar{\boldsymbol{x}}_{i,r}^*$: Again we use the submodularity of valuations to obtain a lower bound on the marginal valuation of $\bar{\boldsymbol{x}}_{i,r+1}^*$.

ditto other def.

$$\begin{aligned} v_i(\bar{\boldsymbol{x}}_{i,r+1}^* \mid h_i^1, \dots, h_i^r) &= v_i(\bar{\boldsymbol{x}}_{i,r+1}^* \cup \{h_i^r\} \mid h_i^1, \dots, h_i^{r-1}) - v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) && (15)\\ &= v_i(\bar{\boldsymbol{x}}_{i,r}^* \setminus \mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^{r-1}) - v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) && (16)\\ &\geq v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1}) - v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) && (17)\\ &\qquad - v_i(\mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^{r-1}) \end{aligned}$$

Case 2 — $h_i^r \notin \bar{\boldsymbol{x}}_{i,r}^*$: We use the monotonicity of valuations for the inequality

$$\begin{aligned}
v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^r) &= v_i(\bar{\boldsymbol{x}}_{i,r}^* \cup \{h_i^1, \dots, h_i^r\}) - v_i(h_i^1, \dots, h_i^r) && (18)\\
&\geq v_i(\bar{\boldsymbol{x}}_{i,r}^* \cup \{h_i^1, \dots, h_i^{r-1}\}) - v_i(h_i^1, \dots, h_i^r) && (19)\\
&= \left(v_i(\bar{\boldsymbol{x}}_{i,r}^* \cup \{h_i^1, \dots, h_i^{r-1}\}) - v_i(h_i^1, \dots, h_i^{r-1})\right) && (20)\\
&\qquad - \left(v_i(h_i^1, \dots, h_i^r) - v_i(h_i^1, \dots, h_i^{r-1})\right)\\
&= v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1}) - v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) && (21)
\end{aligned}$$

after first using the submodularity twice to obtain the same lower bound again:

rethink formulation

$$\begin{aligned}
v_i(\bar{\boldsymbol{x}}_{i,r+1}^* \mid h_i^1, \dots, h_i^r) &= v_i(\bar{\boldsymbol{x}}_{i,r}^* \setminus \mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^r) && (22)\\
&\geq v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^r) - v_i(\mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^r) && (23)\\
&\geq v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^r) - v_i(\mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^{r-1}) && (24)\\
&\geq v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1}) - v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) && (25)\\
&\qquad - v_i(\mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^{r-1})
\end{aligned}$$

In both cases, we can replace $v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1})$ by the induction hypothesis and $v_i(\mathcal{L}_{i,r+1} \mid h_i^1, \dots, h_i^{r-1})$ by $\ell_{i,r+1} \cdot v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1})$ to prove the lemma. For a more detailed formula simplification, we refer to Garg, Kulkarni and Kulkarni [1, p. 14]. □

i: I hope the shortcut is allowed. Might remove it if rest of document not too long.

The lemma can be used to bound the marginal valuation of the items assigned in each round r.

Corollary 1. *From lemma 3 follows*

$$v_i(h_i^r \mid h_i^1, \dots, h_i^{r-1}) \geq \left(u_i^* - \ell_{i,2} v_i(h_i^1) - \sum_{r'=2}^{r-1} \ell_{i,r'+1} \cdot v_i(h_i^{r'} \mid h_i^1, \dots, h_i^{r'-1}) - v_i(h_i^1, \dots, h_i^{r-1})\right) \Big/ \left(\bar{\tau}_{i,0}^* - \sum_{l=1}^{r} \ell_{i,l}\right)$$

where $\bar{\tau}_{i,0}^* := |\bar{\boldsymbol{x}}_{i,0}^*|$ denotes the number of optimal and attainable items after phase I.

Proof. There are $\bar{\tau}_{i,0}^*$ optimal and attainable items in $\bar{\boldsymbol{x}}_{i,0}^*$ at the start of phase II. Of those, $\ell_{i,l}$ many are assigned to other agents in each round $l \leq r$, and also some items h_i^l assigned to agent i may be optimal, whence an upper bound of $\bar{\tau}_{i,0}^* - \sum_{l=1}^{r} \ell_{i,l}$ on the number $\bar{\tau}_{i,r}^*$ of items in the set $\bar{\boldsymbol{x}}_{i,r}^*$.

Possibly this whole estimation can be omitted as we ditch it in the following lemma.

The valuations are monotonic, i. e., $v_i(\mathcal{S}_1) \leq v_i(\mathcal{S}_2)$ for all sets $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \mathcal{G}$ of items. Induction shows that there must be an item $j \in \bar{\boldsymbol{x}}_{i,r}^*$ with a marginal valuation of at least $v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1})/\bar{\tau}_{i,r}^*$. As item h_i^r was the one to be assigned, the marginal valuation of it cannot be smaller. Using lemma 3 for the value of $v_i(\bar{\boldsymbol{x}}_{i,r}^* \mid h_i^1, \dots, h_i^{r-1})$ proves the corollary. □

This, finally, enables us to give a lower bound on the valuation of the whole bundle assigned in phase II.

Lemma 4. *For each agent $i \in \mathcal{A}$ and her bundle $\boldsymbol{x}_i^{\mathrm{II}} = \{h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}}\}$ assigned in phase II, her valuation $v_i(\boldsymbol{x}_i^{\mathrm{II}})$ of her bundle is lower-bounded by her valuation $u_i^* = v_i(\bar{\boldsymbol{x}}_{i,1}^*)$ of the optimal and attainable items in $\bar{\boldsymbol{x}}_{i,1}^*$ after the first round divided by the number n of agents, i. e.,*

$$v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}}) \geq u_i^*/n.$$

Proof. In each round $r = 1, \dots, \tau_i^{\mathrm{II}}$, $\ell_{i,r}$ optimal and attainable items of agent i are assigned to other agents. As there are n agents in total, $n-1$ is an upper bound on $\ell_{i,r}$. Furthermore, after τ_i^{II} rounds, the number $\bar{\tau}^*_{i,\tau_i^{\mathrm{II}}} \leq \bar{\tau}^*_{i,0} - \sum_{l=1}^{r} \ell_{i,l}$ of optimal and attainable items is at most $n-1 \leq n$ else she would have been assigned yet another item. Together with corollary 1, this proves the lemma:

i: swapping the summands' positions reduces the number of lines used

$$
\begin{aligned}
v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}}) &= v_i(h_i^{\tau_i^{\mathrm{II}}} \mid h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}}) + v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) && (26)\\
&\geq \Big(u_i^* - \ell_{i,2} v_i(h_i^1) - \sum_{r'=2}^{\tau_i^{\mathrm{II}}-1} \ell_{i,r'+1} \cdot v_i(h_i^{r'} \mid h_i^1, \dots, h_i^{r'-1}) && (27)\\
&\qquad - v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \Big) \Big/ \Big(\bar{\tau}^*_{i,0} - \sum_{l=1}^{\tau_i^{\mathrm{II}}} \ell_{i,l} \Big) + v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \\
&\geq \Big(u_i^* - (n-1) v_i(h_i^1) - \sum_{r'=2}^{\tau_i^{\mathrm{II}}-1} (n-1) \cdot v_i(h_i^{r'} \mid h_i^1, \dots, h_i^{r'-1}) && (28)\\
&\qquad - v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \Big) \Big/ n + v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \\
&\geq \big(u_i^* - (n-1) v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) - v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \big) \big/ n && (29)\\
&\qquad + v_i(h_i^1, \dots, h_i^{\tau_i^{\mathrm{II}}-1}) \\
&= u_i^*/n && (30)
\end{aligned}
$$

□

Remark 2. possibly not or shorter

Lemma 5. *about phase I & III*

Theorem 2. *RepReMatch has an approximation factor of* $2n(\log n + 2)$.

Proof. □

Remark 3. possibly not or shorter

Algorithm 2: RepReMatch for the Asymmetric Submodular NSW problem

Input : set $\mathcal{A} = \{1, \dots, n\}$ of agents with weights $\eta_i \forall i \in \mathcal{A}$, set $\mathcal{G} = \{1, \dots, m\}$ of indivisible items, submodular valuations $v_i \colon \mathcal{P}(\mathcal{G}) \to \mathbb{R}_0^+$ where $v_i(\mathcal{S})$ is the valuation of agent $i \in \mathcal{A}$ for each set $\mathcal{S} \subset \mathcal{G}$ of items

Output: $\frac{1}{2n(\log n+2)}$-approximation $\boldsymbol{x}^{\mathrm{III}} = (\boldsymbol{x}_1^{\mathrm{III}}, \dots, \boldsymbol{x}_n^{\mathrm{III}})$ of an optimal allocation

Phase I:

$\boldsymbol{x}_i^{\mathrm{I}} \leftarrow \emptyset \quad \forall i \in \mathcal{A}$

$\mathcal{G}^{\mathrm{rem}} \leftarrow \mathcal{G}$

for $t \leftarrow 0, \dots, \lceil \log n \rceil - 1$ **do**

  if $\mathcal{G}^{\mathrm{rem}} \neq \emptyset$ **then**

    $\mathcal{W} \leftarrow \{\, \eta_i \cdot \log(v_i(j)) \mid i \in \mathcal{A}, j \in \mathcal{G} \,\}$

    $G \leftarrow (\mathcal{A}, \mathcal{G}, \mathcal{W})$

    $\mathcal{M} \leftarrow \text{max\_weight\_matching}(G)$

    $\boldsymbol{x}_i^{\mathrm{I}} \leftarrow \boldsymbol{x}_i^{\mathrm{I}} \cup \{j\} \quad \forall (i,j) \in \mathcal{M}$

    $\mathcal{G}^{\mathrm{rem}} \leftarrow \mathcal{G}^{\mathrm{rem}} \setminus \{\, j \mid (i,j) \in \mathcal{M} \,\}$

  end if

end for

Phase II:

$\boldsymbol{x}_i^{\mathrm{II}} \leftarrow \emptyset \quad \forall i \in \mathcal{A}$ ▷ *put allocation $\boldsymbol{x}^{\mathrm{I}}$ aside and start a new one*

while $\mathcal{G}^{\mathrm{rem}} \neq \emptyset$ **do**

  $\mathcal{W} \leftarrow \{\, \eta_i \cdot \log(v_i(\boldsymbol{x}_i^{\mathrm{II}} \cup \{j\})) \mid i \in \mathcal{A}, j \in \mathcal{G} \,\}$

  $G \leftarrow (\mathcal{A}, \mathcal{G}, \mathcal{W})$

  $\mathcal{M} \leftarrow \text{max\_weight\_matching}(G)$

  $\boldsymbol{x}_i^{\mathrm{II}} \leftarrow \boldsymbol{x}_i^{\mathrm{II}} \cup \{j\} \quad \forall (i,j) \in \mathcal{M}$

  $\mathcal{G}^{\mathrm{rem}} \leftarrow \mathcal{G}^{\mathrm{rem}} \setminus \{\, j \mid (i,j) \in \mathcal{M} \,\}$

end while

Phase III:

$\mathcal{G}^{\mathrm{rem}} \leftarrow \bigcup_{i \in \mathcal{A}} \boldsymbol{x}_i^{\mathrm{I}}$ ▷ *release items assigned in phase I*

$\mathcal{W} \leftarrow \{\, \eta_i \cdot \log(v_i(\boldsymbol{x}_i^{\mathrm{II}} \cup \{j\})) \mid i \in \mathcal{A}, j \in \mathcal{G} \,\}$

$G \leftarrow (\mathcal{A}, \mathcal{G}, \mathcal{W})$

$\mathcal{M} \leftarrow \text{max\_weight\_matching}(G)$

$\boldsymbol{x}_i^{\mathrm{III}} \leftarrow \boldsymbol{x}_i^{\mathrm{II}} \cup \{j\} \quad \forall (i,j) \in \mathcal{M}$

$\mathcal{G}^{\mathrm{rem}} \leftarrow \mathcal{G}^{\mathrm{rem}} \setminus \{\, j \mid (i,j) \in \mathcal{M} \,\}$

$\boldsymbol{x}_i^{\mathrm{III}} \leftarrow \text{arbitrary\_allocation}(\mathcal{A}, \mathcal{G}^{\mathrm{rem}}, \boldsymbol{x}_i^{\mathrm{III}}, (v_i)_{i \in \mathcal{A}})$

return $\boldsymbol{x}^{\mathrm{III}}$

decide on $\boldsymbol{x}$ or $\boldsymbol{x}^{\mathrm{III}}$; discrepancies in orig paper in def of $\boldsymbol{x}^{\mathrm{III}}$!