Sequential and Swap Mechanisms for Public Housing Allocation with Quotas and Neighbourhood-based Utilities

Nathanaël Gross–Humbert, Nawal Benabbou, Aurélie Beynier, Nicolas Maudet

Sequential and Swap Mechanisms for Public Housing Allocation with Quotas and Neighbourhood-based Utilities

NATHANAËL GROSS–HUMBERT, NAWAL BENABBOU,
AURÉLIE BEYNIER, and NICOLAS MAUDET, LIP6 - CNRS, Sorbonne Université, France

We consider the problem of allocating indivisible items to agents where both agents and items are partitioned into disjoint groups. Following previous works on public housing allocation, each item (or house) belongs to a block (or building) and each agent is assigned a type (e.g. ethnicity group). The allocation problem consists in assigning at most one item to each agent in a *good* way while respecting diversity constraints. Based on Schelling's seminal work, we introduce a generic individual utility function where the welfare of an agent not only relies on her preferences over the items but also takes into account the fraction of agents of her own type in her own block. In this context, we investigate the issue of stability, understood here as the absence of mutually improving swaps, and we define the cost of requiring it. Then we study the behaviour of two existing allocation mechanisms: an adaptation of the sequential mechanism used in Singapore and a distributed procedure based on mutually improving swaps of items. We first present the theoretical properties of these two allocation mechanisms and we then compare their performances in practice through an experimental study.

CCS Concepts: • **Computing methodologies → Artificial intelligence**.

Additional Key Words and Phrases: Multiagent Resource Allocation, Diversity Constraints, Distributed Allocation Mechanisms, Computational Social Choice.

ACM Reference Format:
Nathanaël Gross–Humbert, Nawal Benabbou,, Aurélie Beynier, and Nicolas Maudet. 2022. Sequential and Swap Mechanisms for Public Housing Allocation with Quotas and Neighbourhood-based Utilities. 1, 1 (October 2022), 24 pages.

1 INTRODUCTION

Fairly dividing indivisible items among agents is a central problem in multiagent systems (see [8], and [5, 24] for recent surveys). There are often relations connecting both items (*e.g.* spatial or temporal relations [8]) and agents (*e.g.* belonging to the same hierarchical structure, or being of the same type). In public housing allocation problems for instance, agents get assigned to locations (houses), belonging to blocks (or buildings). They may of course have preferences over those locations, but importantly, this is also a setting where externalities naturally occur: it makes a difference whether your friends, for instance, get assigned to the same block as you. While agents may naturally seek the proximity of other agents of the same type (a phenomenon well-known as *homophily*), the objective might be opposite at the society level. From the designer's perspective, it is indeed often desirable to preserve some diversity. In practice this can be done by imposing some quotas. Recently, several papers have studied variants of these settings (we review the most relevant literature below). However to the best of our knowledge, none of them addressed a model where agents are motivated (to some extent) by such an homophily bias, while the system has a conflicting diversity objective enforced through a system of quotas. In this paper we undertake the study of such a model. Our main research question is to understand the interplay

Authors' address: Nathanaël Gross–Humbert, Nathanael.Gross--Humbert@lip6.fr; Nawal Benabbou, nawal.benabbou@lip6.fr;
Aurélie Beynier, aurelie.beynier@lip6.fr; Nicolas Maudet, nicolas.maudet@lip6.fr, LIP6 - CNRS, Sorbonne Université, 4 place Jussieu, F-75005 Paris, France.

between these notions, and to investigate the behaviour of some existing allocation mechanisms. We place ourselves in the setting of public housing allocation, and compare (an adaptation of) the existing sequential mechanism used in Singapore, with a simple swap dynamics whereby agents can exchange items when this is mutually beneficial for them, until a stable allocation is reached. Stability has been identified as a key feature in market design [22], evidenced by a number of non-stable mechanisms abandoned in recent history. Specifically, the basic notion of swap stability we study in this paper is of high practical relevance because the opportunity of an improving exchange is likely to be noticed by agents (unlike with the stronger notion of Pareto-optimality). This may encourage some of them to perform deals outside of the mechanism, thus jeopardising its long-term survival, and raising issues of fairness. While the sequential mechanism used in Singapore is appealing because of its simplicity, we will see that it fails to guarantee swap stability. On the other hand, we will show that the above mentioned swap dynamics is guaranteed to terminate even when the utility function accounts for externalities. Our paper explores both theoretically and experimentally how these mechanisms compare, especially in terms of social welfare, and how combining them might be a viable solution in practice.

1.1 Related work

The seminal work of Schelling [23] describes an agent-based model that explains the emergence of global residential segregation in metropolitan areas. In the Schelling model, agents are divided into two groups (or types) that might represent ethnic origins, social characteristics or economic status. The satisfaction of an agent depends on the fraction of agents of her own type in her neighbourhood. If this fraction is below a given tolerance threshold τ, then the agent is unsatisfied. An unsatisfied agent can move to another unoccupied location or can swap with another unsatisfied agent. Schelling demonstrated that the system converges to large regions of homogeneous agents, i.e. regions composed of agents of the same type. This phenomenon is observed even when τ is small.

Chauhan et al. [11] enriched Schelling's model by allowing the agents to have preferences over locations. While in Schelling's model the new location of an agent is chosen at random among satisfying locations, Chauhan et al. considered strategic agents with two goals. The first goal of an agent is to find a location that exceeds her tolerance threshold. Among the set of such locations, an agent then tries to be as close as possible to her favourite location. Elkind et al. [13] extended these models to k types of agents and introduced social Schelling games where possible locations are represented by an undirected graph. Agents are assumed to care only about their location independently of their neighbourhood or to care only about their surroundings independently of their location. Elkind et al. studied the existence of equilibria on different graph topologies and they investigated both the Price of Anarchy (PoA) and the Price of Stability (PoS).

In a different paper, Elkind et al. [14] studied the problem of allocating plots of land to agents that have preferences over the plots and over the number of friends in their neighbourhood. Neighbouring plots are represented as an undirected graph. Elkind et al. proved that even when all agents assign the same value to all their friends, the problem of maximizing the social welfare is NP-hard.

Utility functions based on the homogeneity of the surroundings have also been studied in Fractional Hedonic Games [4, 9, 20]. In such games, each agent assigns a real value to every player and aims to join the coalition which maximizes the sum of her utilities for the agents of the coalition divided by the size of the coalition. In Simple Fractional Hedonic Games, all the values assigned to the other agents are either 1 or 0. Such games allow for formalizing social cohesive groups [4, 18] where each agent wants to maximize the fraction of agents of her own type in the coalition.

Finally, requiring diversity constraints has been recently studied in various two-sided matching settings such as the course allocation problem, the hospital resident problem and college admissions [26]. Most existing approaches consider the point of view of mechanism design and propose matching mechanisms guaranteeing desired properties such as fairness

[10] or incentive incompatibility [3]. Ágoston et al. [26] proposed integer programming formulations for two-sided matchings with diversity constraints and studied the stability of the solutions.

Benabbou et al. [6] introduced the public-housing allocation problem under diversity constraints inspired from the Singaporean public housing system. Each agent has a type (or group) and each item belongs to a block. The allocation of items to agents has to respect block diversity constraints limiting the percentage of agents of a type in the same block. The framework is very close to the one studied in this paper but the utility function studied in [6] only considered the utility for the items (i.e. called *the item-based utility* in this paper). The authors describe the sequential mechanism used to allocate new apartments to residents of Singapore and study how diversity constraints impact the social welfare.

1.2 Outline of the paper

This paper is organized as follows: in Section 2, we provide a formal definition of the allocation problem under study and introduce our general utility model. In Section 3, we study the price of stability, and we show (among other things) that it is upper bounded by 2 in the general case. In Section 4, we focus on a sequential mechanism that is used to solve real public housing problems. In particular, we show that it does not always return a stable allocation, and that its worst-case error (or utility loss) is unbounded in the general case; a tight upper bound is obtained for a special case. In Section 5, we consider a distributed allocation mechanism based on mutually improving swaps of items, and we show that it always reaches a stable outcome after a finite number of steps. We also show that its worst-case error is unbounded in the general case, and a tight upper bound is obtained for a special case. We conclude with some experiments (Section 6) studying the stability and the efficiency of the existing sequential mechanism, and of the swap-deal mechanism which may possibly take place after the sequential allocation.

2 OUR MODEL

In this paper, we consider an allocation problem involving a set $\mathcal{N}$ of n agents, partitioned into a set T of k types $T_1, \ldots, T_k$, and a set $\mathcal{M}$ of m items/houses, partitioned into a set B of l blocks $B_1, \ldots, B_l$, where the inequality $|\mathcal{N}| \geq |\mathcal{M}|$ holds; note that it is a realistic assumption, especially when considering the allocation of public goods. We denote by $\mathcal{T}(i)$ the type of any agent $i \in \mathcal{N}$ and by $\mathcal{B}(h)$ the block of any item $h \in \mathcal{M}$. Following the work of Benabbou et al. [6], diversity constraints are here defined using type-block capacities/quotas $\lambda_{p,q} \in \mathbb{N}$, with $(p, q) \in [k] \times [l]$, such that $\lambda_{p,q}$ stands for the maximum number of agents of type T_p allowed in block B_q. Without loss of generality, we assume that the inequality $\lambda_{p,q} \leq |B_q|$ holds for all $(p, q) \in [k] \times [l]$ since it is not possible to assign more than $|B_q|$ items in block B_q by definition. We also assume that the inequality $\sum_{p \in [k]} \lambda_{p,q} \geq |B_q|$ holds for all blocks $q \in [l]$ otherwise all allocations satisfying diversity constraints would leave some items unassigned.

Definition 1 (Valid Allocation). *An allocation $A : \mathcal{N} \rightarrow 2^{\mathcal{M}}$ is a function that maps every agent $i \in \mathcal{N}$ to a subset $A(i) \subset \mathcal{M}$ of items. An allocation A is* valid *iff:*

(1) $\forall i \in \mathcal{N}, |A(i)| \leq 1$ *(each agent gets at most one item).*
(2) $\forall i, j \in \mathcal{N},\ A(i) \cap A(j) = \emptyset$ *(agents do not share items).*
(3) $\bigcup_{i \in \mathcal{N}} A(i) = \mathcal{M}$ *(all items are assigned).*
(4) $\forall (p, q) \in [k] \times [l],\ |\{i \in T_p : A(i) \in B_q\}| \leq \lambda_{p,q}$ *(diversity constraints are satisfied).*

Note that checking whether there exists a valid allocation or not can be performed in polynomial time using a network flow formulation [6].

Utility model. We assume here that the utility $u_i(A)$ of an agent $i \in \mathcal{N}$ for an allocation A has two components:

- $u_i^I(A) \in [0,1]$: an *item-based utility* representing the utility derived by agent i for the item $A(i)$ she receives. If i does not receive any item (i.e. $A(i) = \emptyset$), then $u_i^I(A) = 0$. For simplicity, we will denote by $u_i^I(h)$ the item-based utility of agent i for any item $h \in \mathcal{M}$.
- $u_i^N(A) \in [0,1] \cap \mathbb{Q}$: a *neighbour-based utility* which is equal to the fraction of agents of type $\mathcal{T}(i)$ assigned to items in block $\mathcal{B}(A(i))$. More formally, when i gets an item, it is defined by:
$$u_i^N(A) = \frac{\sum_{j \in \mathcal{N}: A(j) \in \mathcal{B}(A(i))} \mathbb{I}(\mathcal{T}(i), \mathcal{T}(j))}{|\mathcal{B}(A(i))|}$$
where $\mathbb{I}(\mathcal{T}(i), \mathcal{T}(j))$ equals 1 if agents i and j have the same type, and equals 0 otherwise. If the agent i gets no item, then $u_i^N(A) = 0$.

The utility of agent $i \in \mathcal{N}$ for allocation A is defined by:

$$u_i(A) = u_i^I(A) + \varphi_i \times u_i^N(A)$$

where $\varphi_i \in [0,1]$ is called the *utility trade-off* and is used to define the relative importance of the item-based utility and the neighbour-based utility. This type of utility function thus allows to model agents which are both concerned by the item they obtain, as well as their neighbourhood. In this paper, a special focus will be placed on the following two types of behaviour: *item-focused* and *neighbour-focused* agents. Agents are said to be *item-focused* if they only care about the item they receive. Agents are said to be *neighbour-focused* when they only care about the neighbourhood. More formally:

DEFINITION 2 (ITEM-FOCUSED AGENTS). *Agent $i \in \mathcal{N}$ is said to be* item-focused *when $\varphi_i = 0$ and $u^I(A) \neq 0$ for some valid allocation A.*

DEFINITION 3 (NEIGHBOUR-FOCUSED AGENTS). *Agent $i \in \mathcal{N}$ is said to be* neighbour-focused *when $\varphi_i \neq 0$ and $u^I(A) = 0$ for every valid allocation A.*

Social welfare. When assessing the welfare of the whole society of agents, we rely on the classical utilitarian social welfare:
$$sw(A) = \sum_{i \in \mathcal{N}} u_i(A)$$

The house allocation problem with quotas and neighbourhood-based utilities. An instance $\mathcal{I}$ of this problem is a tuple $\mathcal{I} = \langle \mathcal{N}, \mathcal{M}, B, T, u^I, u^N, \lambda, \varphi \rangle$ with:

- $\mathcal{N}, \mathcal{M}, B, T$ as defined above,
- $\lambda = \langle \lambda_{1,1}, \cdots, \lambda_{k,l} \rangle$ the $[k] \times [l]$ matrix of quotas,
- $u^I = \langle u_1^I, \cdots, u_n^I \rangle$ the item-based utility profile,
- $u^N = \langle u_1^N, \cdots, u_n^N \rangle$ the neighbour-based utility profile,
- $\varphi = \langle \varphi_1, \cdots, \varphi_n \rangle$ the utility trade-offs.

Note that the problem of maximizing the utilitarian social welfare is computationally difficult since it is NP-complete even in the special case where $\varphi_i = 0$ for all $i \in \mathcal{N}$ (as proved in [6]). It is also known that the problem of maximizing the utilitarian social welfare is NP-complete even when the number of agents of each type is at most 2 and there is only one block with 2 items while all the other blocks contain 1 item (translation of Theorem 3.2 from [14] to our context). In Appendix A, we propose two mixed-integer linear programs, involving a polynomial number of variables and constraints, for the computation of a valid allocation maximizing the utilitarian social welfare and a stable valid allocation of maximum utilitarian social welfare. In this paper, we will use the same utility trade-off for all agents (simply denoted by φ), as a first step in the study of the interplay between our neighbourhood-based utility model and type-block quotas. The general case considering possibly different utility trade-offs is discussed in Appendix F.

Motivation and relation to the state of the art. Our utility model takes inspiration from the model of [14] but the differences are important to notice. First, both our type and neighbourhood relations are partitions among agents ("being of the same type" or "living in the same block"), while they are arbitrary (undirected) graphs in [14]: they would thus correspond to collections of cliques in their model. [17] studied the dorm assignment problem where a resource is shared by several agents. They considered a similar utility function to ours where the utility of an agent i is defined from the utility for the resource and but also from externalities depending on the other agents sharing the resource with the agent i. In this context, [17] have been interested in Pareto envy-free assignments.

Now regarding the neighbour-based utility specifically, by taking the ratio of agents of the same type, our model is closer in spirit to the model of Simple Symmetric Fractional Hedonic Games (SSFHG) [4] where each block can be viewed as a coalition. In such a SSFHG, each agent would assign 1 to an agent of her own type in her own block and 0 to the other agents. The game would be symmetric since for all couples of agents, both agents assign the same value to each other. However, in our context, the size of the blocks (i.e. the coalitions in a SSFHG) is fixed beforehand. This is an important difference that distinguishes our work from Fractional Hedonic Games. To compare to [14] again, our setting would correspond to the *binary friendship-uniform and symmetric* case (instances where the preferences of an agent for the others are either 0 or 1, and they are symmetric). Unlike [14], our neighbour-based utility averages the values for the other agents in the block (i.e. neighbourhood) over the size of the block. This definition is similar to Fractional Hedonic Games but differs from the friendship utility introduced in [14] except when all blocks have the same size.

Finally, we note that our definition implicitly assumes that an agent counts herself among the agents of her type. To motivate our choice, consider the situation where the residents of a block have to make decisions locally to their block. When decisions are taken from a voting rule (the majority rule for instance), an agent would want to form the largest coalition in the block with agents sharing her opinion or her preferences on possible decisions. Moreover, with this definition, the model remains consistent when an agent is alone: an agent would prefer to be alone in a small block than in a large block where it would fear to be ignored. Note that not counting the agent herself would lead to a utility of 0 for blocks with different sizes leading the agent to be indifferent between these blocks.

We conclude the section with a technical lemma on the largest utility ratio which may occur between any two valid allocations – it will be useful when analyzing the worst-case utility losses of allocation procedures.

LEMMA 1. *The largest utility ratio between any two valid allocations is unbounded in the general setting. When $\varphi \neq 0$ and k is a constant, it is upper bounded by $\frac{1+\varphi}{\varphi}k$ and the bound is tight.*

The proof is given in Appendix B. Thus the largest utility ratio between any two valid allocations grows with k and inversely with φ. For instance, in the Singapore public housing system, agents are partitioned into $k = 3$ types (ethnic groups), and therefore we already know that the worst-cast error of any house allocation procedure returning a valid allocation would be upper bounded by 6 when item-based and neighbour-based utilities are given the same importance (i.e. $\varphi = 1$), whereas it would be upper bounded by 9 when the former is twice as important as the latter (i.e. $\varphi = 0.5$).

3 THE PRICE OF STABILITY

A key property of an allocation is *stability*, whether it is such that no agent would like to deviate from the prescribed allocation. Roth suggested that stability is an important criterion to develop a successful matching mechanism [22]. In our context, we shall concentrate on the notion of *swap-stability* [2, 19]: it shouldn't be the case that two agents would be happy to swap their items (i.e perform a swap-deal), resulting in a valid allocation.

DEFINITION 4 (IMPROVING SWAP-DEAL). *Given a valid allocation A, a swap-deal among a pair of agents $(i, j) \in \mathcal{N} \times \mathcal{N}$ is said to be* improving *if and only if $u_i(A_{i\Leftrightarrow j}) > u_i(A)$ and $u_j(A_{i\Leftrightarrow j}) > u_j(A)$, where $A_{i\Leftrightarrow j}$ is the allocation obtained from A by swapping the items of agents i and j.*

Note that agents are assumed to be myopic and swap-deals only consider the immediate utility improvement. From a given valid allocation A, some improving swap-deals may lead to an invalid allocation due to type-block quotas. We thus restrict the set of swap-deals that can be applied from an allocation as follows.

DEFINITION 5 (VALID SWAP-DEAL). *Given a valid allocation, a swap-deal among a pair of agents $(i, j) \in \mathcal{N} \times \mathcal{N}$ is* valid *if and only if the resulting allocation satisfies the diversity constraints.*

We can now introduce our stability notion and the price of stability.

DEFINITION 6 (STABLE ALLOCATION). *An allocation A is* stable *if and only if there is no valid improving swap-deal from A.*

The price of stability (PoS) is the largest utility ratio between any valid allocation of maximum social welfare and any valid stable allocation of maximum social welfare. More formally:

DEFINITION 7 (PRICE OF STABILITY (POS)). *The Price of Stability (PoS) is defined by:*

$$PoS = \sup_{\mathcal{I}} \frac{sw(A^*_{\mathcal{I}})}{sw(A^{\top}_{\mathcal{I}})}$$

where $A^*_{\mathcal{I}}$ (resp. $A^{\top}_{\mathcal{I}}$) is the maximum social welfare allocation that is valid (resp. valid and stable) for the instance $\mathcal{I}$.

We begin our analysis with the following positive result.

PROPOSITION 1. *$PoS = 1$ when all agents are item-focused. The same applies to neighbour-focused agents.*

PROOF. Here we only need to prove that any valid allocation that maximizes the utilitarian social welfare is stable. For the case of item-focused agents, the argument is straightforward: if two agents wish to swap and they are allowed to do so, then implementing the swap would strictly increase the utilities of both agents, while leaving the other utilities unchanged, which yields an allocation with a strictly larger utility. Therefore any valid allocation maximizing the social welfare is necessarily stable. For the case of neighbour-focused agents, it can be proved that if two agents wish to swap to increase their utilities, the utilitarian social welfare will increase by twice as much (the detailed proof is given in Appendix C). Thus any improving swap-deal would strictly increase the social welfare, which implies that any valid allocation maximizing the social welfare is necessarily stable. □

Unfortunately, this result no longer holds when we consider the general form of our utility model. In that case, it can be that two agents have an incentive to swap their items, but that the externalities on other agents make it overall damaging for social welfare, as shown in the following example.

EXAMPLE 1. *Consider an instance of the problem with a set of* 6 *agents* $\mathcal{N} = \{1, 2, 3, 4, 5, 6\}$ *partitioned into* 2 *types* $T_1 = \{1, 2, 3\}$, $T_2 = \{4, 5, 6\}$, *and a set of* 6 *items* $\mathcal{M} = \{h_1, h_2, h_3, h_4, h_5, h_6\}$ *partitioned into* 2 *blocks* $B_1 = \{h_1, h_2, h_3\}$ *and* $B_2 = \{h_4, h_5, h_6\}$. *Assume that type-block quotas are defined by* $\lambda_{p,q} = 3$ *for all* $(p, q) \in \{1, 2\} \times \{1, 2\}$. *Now set* $\varphi = 1$ *and define the item-based utilities as follows:* $u^I_1(h_6) = \frac{3}{4}$, $u^I_6(h_1) = \frac{3}{4}$, $u^I_i(h_i) = 1$ *for all* $i \in \{2, 3, 4, 5\}$, *and* $u^I_i(h_j) = 0$ *for any other agent-item pair* (i, j). *In the optimal allocation, agents* $1, 2$ *and* 3 *get assigned to* h_1, h_2, *and* h_3 *respectively (in block* B_1*), while agents* $4, 5$, *and* 6 *get assigned to* h_4, h_5, *and* h_6 *respectively (in block* B_2*), leading to a social welfare of* $(0+1)+(1+1)+(1+1)+(1+1)+(1+1)+(0+1) = 10$. *But take agents 1 and 6: by swapping their items, they would gain* $\frac{3}{4}$ *in terms of item-based utility, and lose* $\frac{2}{3}$ *in terms of neighbour-based utility. Therefore, the allocation is not stable. Observe that in the resulting allocation* $A_{1\Leftrightarrow 6}$, *we have* $sw(A_{1\Leftrightarrow 6}) = (\frac{3}{4}+\frac{1}{3})+(1+\frac{2}{3})+(1+\frac{2}{3})+(1+\frac{2}{3})+(1+\frac{2}{3})+(\frac{3}{4}+\frac{1}{3}) = \frac{53}{6} < 10$.

On the positive side, we establish the following lower and upper bounds on the price of stability.

PROPOSITION 2. *In the general case,* $1.6 < PoS \leq 2$.

The arguments of the proof exploit notions introduced in the next sections. For readability reasons, we thus defer the proof of this result in Appendix D.

4 A SEQUENTIAL MECHANISM

In this section, we analyze the sequential procedure presented in [6][1] which is a simplified version of the Singaporean public housing allocation process: in some random order, the agents sequentially pick the items that maximize their utilities at the time of their selection, while respecting the diversity constraints (in other words, they are not allowed to pick items in blocks where type-block quotas are reached). This mechanism, called SEQ-mechanism hereafter, is an example of a (constrained) random serial dictatorship or random priority procedure [1, 7]. This procedure benefits from being simple and requires little information about the agents' preferences: at her turn, an agent only communicates her most preferred item among the remaining admissible items. This is undoubtedly an advantage as discussed in [22]. However, this mechanism may "waste" some items due to diversity constraints (see Proposition 3), leading to invalid allocations. Even when it does return a valid allocation, its worst-case utility loss is unbounded in the general case, and is equal to the largest utility ratio between any two valid allocations when $\varphi \neq 0$ and k is a constant (see Proposition 4). Moreover, it is not guaranteed to return a stable allocation in the general case (see Propositions 5 and 6).

PROPOSITION 3. *SEQ-mechanism does not always return a valid allocation, even when all agents are neighbour-focused.*

PROOF. Consider an instance with a set of 4 neighbour-focused agents $\mathcal{N} = \{1, 2, 3, 4\}$ partitioned into two types $T_1 = \{1, 2\}$ and $T_2 = \{3, 4\}$, a set of 4 items $\mathcal{M} = \{h_1, h_2, h_3, h_4\}$ partitioned into two blocks $B_1 = \{h_1, h_2\}$ and $B_2 = \{h_3, h_4\}$, and the following type-block capacities: $\lambda_{1,1} = \lambda_{2,1} = 1$ (at most one agent per type in block B_1) and $\lambda_{1,2} = \lambda_{2,2} = 2$ (at most two agents per type in block B_2). When SEQ-mechanism is runned with agent order $(1, 2, 3, 4)$, nothing prevent the first two agents from picking the two items available in block B_2, which then forces agent 3 to pick an item in block B_1, leaving agent 4 unassigned since her quota is reached in block B_1. In that case, the resulting allocation is not valid as one item remains unassigned. □

Note that SEQ-mechanism always returns a valid allocation when we further assume that type-quotas satisfy inequality $\sum_{q=1}^{l} \lambda_{p,q} \leq |T_p|$ for all $p \in [k]$. These inequalities are naturally verified in problems where the number of agents is very

[1] This mechanism is referred to as "lottery-based" mechanism in [6].

large compared to the number of items, or when the type-block quotas are somewhat proportional to the percentages of the groups in the general population (see e.g., [6]). Now we analyze the worst-case utility loss of SEQ-mechanism, defined as the largest utility ratio between a social welfare optimum and any valid allocation returned by SEQ-mechanism.

PROPOSITION 4. *The worst-case utility loss of SEQ-mechanism is unbounded in the general case. When $\varphi \neq 0$ and k is a constant, it reaches the upper bound of Lemma 1.*

PROOF. Given a parameter $k \in \mathbb{N}$, we consider a family of instances with k types of agents, k blocks of items and k items per block (i.e. $|B_q| = k$ for all $q \in [l]$). All but the first type of agents are composed of exactly k agents (i.e. $|T_p| = k$ for all $p \in \{2, \dots, k\}$), whereas the first type T_1 is partitioned into two groups T_1^1 and T_1^2 such that $|T_1^1| = k^2$ and $|T_1^2| = k$. Hence we have $n = 2k^2$ agents and $m = k^2$ items. The type-block capacities are as follows: $\lambda_{1,q} = k$ for all $q \in [l]$ and $\lambda_{p,q} = 1$ for all $p \in \{2, \dots, k\}$ and all $q \in [l]$. Finally, we assume that all agents $i \in \mathcal{N} \backslash T_1^1$ are neighbour-focused (i.e., $u_i^I(A) = 0$ for any valid allocation A), while the remaining agents have a utility of 1 for all items (i.e., $u_i^I(A) = 1$ for any valid allocation A). Obviously, any allocation A^* assigning all the items to the agents in T_1^1 maximizes the utilitarian social welfare and we have:

$$sw(A^*) = \sum_{i \in T_1^1} \left(u_i^I(A^*) + \varphi u_i^N(A^*) \right) = \sum_{i \in T_1^1} \left(1 + \varphi \times 1 \right) = k^2 \times (1 + \varphi)$$

Now consider some agent order such that all the agents in $\mathcal{N} \backslash T_1$ are positioned before the agents in T_1^2, which in turn are all positioned before the agents in T_1^1. If SEQ-mechanism is applied with this agent order, then we will obtain a valid allocation A where blocks only include one agent per type, but no agent from T_1^1. The corresponding social welfare is:

$$sw(A) = \sum_{i \in \mathcal{N} \backslash T_1^1} u_i(A) = \sum_{i \in \mathcal{N} \backslash T_1^1} \left(0 + \varphi \times \frac{1}{k} \right) = k^2 \times \varphi \frac{1}{k} = \varphi k$$

Thus the utility loss of SEQ-mechanism is here equal to:

$$\frac{sw(A^*)}{sw(A)} = \frac{k^2 \times (1 + \varphi)}{\varphi k} = \frac{(1 + \varphi)}{\varphi} k$$

which is exactly equal to the bound of Lemma 1. Note that it grows linearly with k, which means that it is unbounded in the general case. □

We now focus on stability.

PROPOSITION 5. *For some instances, SEQ-mechanism never returns a stable valid allocation, whatever the agent order (even when there exist valid stable allocations).*

PROOF. Consider an instance with a set of 4 agents $\mathcal{N} = \{1, 2, 3, 4\}$ partitioned into 2 types $T_1 = \{1, 2\}$ and $T_2 = \{3, 4\}$, and a set of 4 items $\mathcal{M} = \{h_1, h_2, h_3, h_4\}$ partitioned into 2 blocks $B_1 = \{h_1, h_2\}$ and $B_2 = \{h_3, h_4\}$. Assume that quotas are defined by $\lambda_{1,1} = 1$, $\lambda_{2,1} = 2$, $\lambda_{1,2} = 2$ and $\lambda_{2,2} = 1$. Now set $\varphi = 1$ and define the item-based utilities as follows: $u_i^I(h_1) = 1$, $u_i^I(h_2) = u_i^I(h_3) = u_i^I(h_4) = 0.75$ for all agents $i \in T_1$, and $u_i^I(h_3) = 1$, $u_i^I(h_1) = u_i^I(h_2) = u_i^I(h_4) = 0.75$ for all agents $i \in T_2$. Whatever the agent order, the sequential mechanism always returns an allocation with only one agent per type in each block since the first agent always picks the item for which she has an item-based utility of 1, which is in the block where her type-block quota is equal to 1. Such an allocation is not stable as the swap-deal among (i, j) with $i \in T_1$ and $j \in T_2$ such that $A(i) \in B_2$ and $A(j) \in B_1$ is a valid improving swap-deal. Note that it is easy to check that the resulting allocation is a valid stable allocation. □

Proposition 6. *SEQ-mechanism does not always return a stable allocation, even when all agents are neighbour-focused. It always returns a stable allocation when all agents are item-focused.*

Proof. Consider an instance with a set of 4 neighbour-focused agents $\mathcal{N} = \{1, 2, 3, 4\}$ partitioned into 2 types $T_1 = \{1, 2\}$ and $T_2 = \{3, 4\}$, and a set of 4 items $\mathcal{M} = \{h_1, h_2, h_3, h_4\}$ partitioned into 2 blocks $B_1 = \{h_1, h_2\}$ and $B_2 = \{h_3, h_4\}$. Assume that type-block quotas are defined by $\lambda_{p,q} = 2$ for all $(p, q) \in \{1, 2\} \times \{1, 2\}$, which implies that they are not restrictive in this example. When we run the sequential procedure with agent order $(1, 3, 2, 4)$, we can obtain an allocation A where agents 1 and 3 get an item in block B_1 and agents 2 and 4 get an item in block B_2. In that case, each agent $i \in \mathcal{N}$ has a utility $u_i(A) = 0 + \varphi \times 1/2$ since the blocks only include one agent per type. Note that agents 2 and 3 would be happy to swap their items as they would get a utility of $0 + \varphi \times 2/2 > u_i(A)$.

When all agents are item-focused, the proof is straightforward: since their utilities do not depend on the choices made by the other agents, an item-focused agent can only envy an agent that picked an item before her, and therefore it cannot be the case that two item-focused agents wish to exchange their items. □

5 A SWAP-DEAL MECHANISM

A natural distributed approach in multiagent resource allocation is to start from a valid allocation and let the agents perform bilateral improving swap-deals until they reach a stable outcome [12, 15, 25]. Swap-deals require very little coordination: the agents do not need to have full knowledge about the utilities and allocation of the other agents, and only two agents have to interact to make a swap. Under such swap mechanisms, the agents have an initial (valid) endowment, which reflects the current state of affairs; for the Singaporean public housing allocation problem, it could be the outcome of the sequential mechanism, as we just proved that it may not be stable (see Propositions 5 and 6). Here we focus on a simple swap-deal mechanism where at each step, pairs of agents meet randomly and perform swap-deals if possible; this mechanism will be called SWAP-mechanism hereafter.

When all agents are item-focused or neighbour-focused, it follows from the proof of Proposition 1 that SWAP-mechanism will converge to a stable outcome (as any improving swap strictly increases the social welfare). In the general case, our example showing that valid improving swap-deals may decrease social welfare suggests that convergence may not occur (see Example 1). Still, we can prove that the SWAP-mechanism converges to a stable outcome.

Proposition 7. *In the general case, SWAP-mechanism will provably reach a stable outcome.*

Proof. It can be proved that the SWAP-mechanism will always reach an equilibrium by analyzing the potential function *pot* defined by:

$$pot(A) = sw(A) + \sum_{i \in \mathcal{N}} u_i^I(A)$$

for any valid allocation A. To do so, consider some *non-stable* allocation A and some valid improving swap-deal between two agents. Let A' denote the allocation obtained after the swap-deal is performed. The variation in social welfare $\Delta U = sw(A') - sw(A)$ can be decomposed as follows:

$$\Delta U = \varphi(\Delta U_1 + \Delta U_2) + \Delta U_3$$

where ΔU_1 (resp. ΔU_2) is the variation in the neighbour-based utilities of the agents who participate (resp. do not participate) in the swap-deal, and ΔU_3 is the variation in the item-based utilities of the two agents participating in the swap-deal (the item-based utilities do not change for the other agents). We know that $\Delta U_1 = \Delta U_2$ (see the detailed proof of Proposition 1 given in the supplementary materials), we can thus determine that:

$$\begin{aligned} pot(A') - pot(A) &= sw(A') - sw(A) + \sum_{i \in \mathcal{N}} u_i^I(A') - \sum_{i \in \mathcal{N}} u_i^I(A) \\ &= \Delta U + \Delta U_3 \\ &= \varphi \Delta U_1 + \varphi \Delta U_2 + \Delta U_3 + \Delta U_3 \\ &= 2(\Delta U_3 + \varphi \Delta U_1) \end{aligned}$$

Since $\Delta U_3 + \varphi \Delta U_1$ is the increase in utility of the two agents swapping, it is strictly positive (by definition of improving swap-deals). Thus, *pot* strictly increases after each swap-deal. Since *pot* is upper bounded by $3|\mathcal{M}|$ and the number of valid allocations is finite, we will reach an equilibrium after a finite number of steps. □

Now we analyze the worst-case loss of SWAP-mechanism, i.e. the largest utility ratio between a social welfare optimum and any (valid) allocation returned by SWAP-mechanism.

PROPOSITION 8. *The worst-case utility loss of SWAP-mechanism is unbounded in the general case. When $\varphi \neq 0$ and k is a constant, it reaches the upper bound of Lemma 1.*

The proof of Proposition 8 is given in Appendix E. Now we analyze the price of Price of Anarchy (PoA) [16], which is here defined as the largest utility ratio between any valid allocation and any valid stable allocation:

DEFINITION 8 (PRICE OF ANARCHY (POA)). *The Price of Anarchy (PoA) is defined by:*

$$PoA = \sup_{\mathcal{I}} \frac{sw(A^*_{\mathcal{I}})}{sw(A^{\perp}_{\mathcal{I}})}$$

where $A^*_{\mathcal{I}}$ (resp $A^{\perp}_{\mathcal{I}}$) is a valid allocation (resp. valid stable allocation) of instance $\mathcal{I}$ which maximizes (resp. minimizes) the social welfare.

In other words, the PoA (resp. PoS) is the largest utility ratio between any social welfare optimum and one of the *worst* (resp. *best*) valid stable allocations. From Lemma 1 and Proposition 8, one can directly derive the following result on the PoA:

COROLLARY 1. *The PoA is unbounded in the general case. When $\varphi \neq 0$ and k is a constant, it is upper bounded by $\frac{1+\varphi}{\varphi}k$ and the upper bound is tight.*

6 EXPERIMENTS

In this section, we report the results of numerical tests aiming to evaluate the practical performances of the following allocation mechanisms:

- RSEQ: in some random order, the agents sequentially select items *at random* while respecting the diversity constraints. This procedure will be used as a baseline, generating random allocations using an approach similar to that of SEQ-mechanism.
- SEQ: SEQ-mechanism, as described in Section 4, where in some random order, the agents sequentially select the items *maximising their utilities* while respecting the diversity constraints.
- RSEQ+SWAP: the two-phase procedure consisting in first running RSEQ and then applying SWAP-mechanism.
- SEQ+SWAP: the two-phase procedure consisting in first running SEQ and then applying SWAP-mechanism.

These four mechanisms are compared in terms of utilitarian social welfare in Sections 6.1 and 6.2. In Section 6.1, we consider random instances involving only 40 houses/agents, so as to be able to compute the utilitarian social welfare optimum and empirical price of stability (using the linear programs given in Appendix A). In Section 6.2, these four mechanisms are compared using a real dataset from Singapore real estate, which involves 1350 houses. In Section 6.3, we compare RSEQ+SWAP and SEQ+SWAP in terms of number of swaps that are needed to reach a stable allocation; we generate synthetic datasets of different sizes to study the impact of the number of items/agents on the number of swaps.

6.1 Empirical price of stability and comparison of allocation mechanisms

In this subsection, we consider small instances of the house allocation problem with quotas and neighbourhood-based utilities, which allows us to compute the social welfare optimum and empirical price of stability (using the linear programs given in Appendix A). More precisely, we generate 30 instances with 40 houses partitioned into $l = 5$ blocks of same size, and 40 agents partitioned into $k = 5$ types of same size. The type-block quotas are set to $\lambda_{p,q} = 2$ for all $p, q \in [5]$. The item-based utilities are uniformly drawn in $[0, 1]$. We consider two utility trade-offs: $\varphi = 0.5$ and $\varphi = 1$.

For these instances, we have computed the empirical price of stability and it turns out to be exactly equal to 1 —in other words, the social welfare optimum is stable in every generated instance. Table 1 reports the averaged utility loss incurred by the four allocation mechanisms, i.e. the utility ratio between the social welfare optimum and the valid allocation returned by the mechanism[2]. In this table, we observe that RSEQ is outperformed by SEQ, RSEQ+SWAP and SEQ+SWAP, which have relatively similar performances on random data. More precisely, their averaged utility losses are relatively close to the empirical price of stability (when compared to that of RSEQ), and very far from the theoretical upper bound given in Lemma 1 (which is equal to $\frac{1+1}{1} \times 5 = 10$ for $\varphi = 1$ and to $\frac{1+0.5}{0.5} \times 5 = 15$ for $\varphi = 0.5$). However, the averaged utility losses of SEQ and SEQ+SWAP increase with φ, whereas RSEQ+SWAP is relatively constant. This is mainly due to the fact that the first agents selecting items under SEQ-mechanism have no vision on their future neighbourhoods, and they are not allowed to switch items afterwards (the loss is reduced when applying SWAP right after).

	RSEQ	SEQ	RSEQ+SWAP	SEQ+SWAP
$\varphi = 0.5$	1.757 ± 0.144	1.073 ± 0.028	1.101 ± 0.023	1.069 ± 0.029
$\varphi = 1$	1.695 ± 0.118	1.084 ± 0.027	1.102 ± 0.026	1.074 ± 0.021

Table 1. Averaged utility losses on random instances.

6.2 Comparison of mechanisms on real data from Singapore real estate

In this part, we use a dataset from Singapore real estate, with $m = 1350$ houses partitioned into $l = 9$ blocks/buildings of size 128, 162, 156, 249, 108, 94, 104, 190, and 159 respectively (see [6] for a complete description of the dataset). The set $\mathcal{N}$ of agents is divided into $k = 3$ types of size $|T_1| = 0.74 \times n$, $|T_2| = 0.13 \times n$, and $|T_3| = 0.13 \times n$ (following the 2010 Singapore census report), where $n = 2700$ is the number of agents applying for a house. The type-block quotas are defined by $\lambda_{1q} = 0.87 \times |B_q|$, $\lambda_{2q} = 0.25 \times |B_q|$, and $\lambda_{3q} = 0.15 \times |B_q|$ for all blocks $q \in [l]$ (as defined by the Housing and Development Board in Singapore). In our tests, we set φ to 1, and for the item-based utilities, we consider the following two utility models (introduced in [6]):

[2] RSEQ and SEQ needed to be run respectively 1.87 ± 1.74 and 1.83 ± 1.29 times on average to obtain a valid allocation.

- The distance-based utility model (Dist(σ^2)): in this model, each agent i has a preferred geographic location p_i, uniformly drawn at random in $[0, 10]^2$. The utility derived by agent i for item h in block B_q is then generated according to a normal distribution: $\mathcal{N}(1/d(p_i, L(B_q)), \sigma^2)$ where $L(B_q)$ is the location of block B_q and $d(\cdot)$ is the Euclidean distance.
- The ethnicity-based utility model (Ethn(σ^2)): in this model, all agents of the same type have the same preferred location.

We consider two values of σ^2 (variance): 0.01 and 1. In Table 2, we report the social welfare[3] of the allocations constructed by the four allocation mechanisms, averaged over 30 runs. In this table, we observe that RSEQ is here also outperformed by SEQ, RSEQ+SWAP, and SEQ+SWAP, in terms of social welfare. The latter three allocation procedures achieve similar performances on all settings, except for Dist(0.01) where SEQ is significantly less efficient than SEQ+SWAP and RSEQ+SWAP. This is due to the fact that agents of the same type have very different item-based utility values under Dist(0.01), which makes them choose different blocks at the first iteration steps of SEQ. Interestingly though, adding SWAP after SEQ allows to significantly improve the social welfare in that case.

	Dist(σ^2)		Ethn(σ^2)	
	$\sigma^2 = 0.01$	$\sigma^2 = 1$	$\sigma^2 = 0.01$	$\sigma^2 = 1$
RSEQ	903 ± 14	1465 ± 43	1133 ± 84	1466 ± 41
SEQ	989 ± 23	1678 ± 60	1264 ± 99	1706 ± 60
RSEQ+SWAP	1057 ± 19	1659 ± 58	1242 ± 93	1684 ± 55
SEQ+SWAP	1072 ± 23	1690 ± 61	1268 ± 101	1716 ± 61

Table 2. Averaged social welfare values on the dataset from Singapore real estate.

6.3 Empirical number of swaps

Now we analyse the number of swaps needed by SWAP-mechanism to reach a stable allocation. More precisely, we compare RSEQ+SWAP and SEQ+SWAP on randomly generated instances with $l = 10$ blocks of equal size, varying the number of agents/items from 100 to 1400 (in steps of 100). The set of agents is divided into $k = 3$ types of size $|T_1| = 0.74 \times n$, $|T_2| = 0.13 \times n$, and $|T_3| = 0.13 \times n$, and type-block quotas are defined by $\lambda_{1q} = 0.87 \times |B_q|$, $\lambda_{2q} = 0.25 \times |B_q|$, and $\lambda_{3q} = 0.15 \times |B_q|$ for all blocks $q \in [l]$ (following the dataset from Singapore real estate). Item-based utilities are uniformly drawn at random within $[0, 1]$ and φ is set to 1.

In Figure 1, we report the results obtained on 20 instances (for every value of n), and linear regressions are applied on the corresponding data points:

- For the blue points (RSEQ+SWAP), we obtain a slope of 1.88, an intercept of -141.20 and a correlation factor of 0.998 (very high correlation);
- For the red points (SEQ+SWAP), we obtain a slope of 0.08, an intercept of 43.53, and a correlation factor of 0.74 (high correlation).

[3] We do not report the averaged utility losses here, as it was not possible to compute the social welfare optimum efficiently using our linear programming formulation (due to the high number of agents/items).

Thus for both procedures, the number of swaps grows linearly with the number of agents in practice. However, SEQ+SWAP is here clearly the best procedure as it requires far fewer swaps than RSEQ+SWAP to reach a stable outcome. In fact, that number grows very moderately as the size of the instance augments.

Other experiences were made with constant numbers of agents, items and types, but varying the number of blocks. The results suggested that the number of blocks has almost no impact on the empirical time complexity of SWAP-mechanisms.

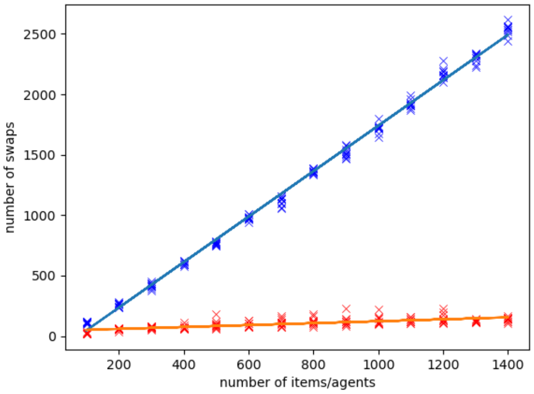

Fig. 1. Linear regression on the number of swaps needed to reach an equilibrium.

7 CONCLUSION

This paper investigated a model where, while agents have an homophily component in their utility function (whose relative importance is determined by a parameter φ), there is a society-wide objective to promote diversity (in that case, through the use of quotas). Our work has connection to (and takes inspiration from) a number of recent works, but it is unique in the combination of these different features. We analyse the properties of two mechanisms: the sequential approach used in Singapore, and the swap-deal approach letting agents exchange their items when this is beneficial. While both exhibit the same worst-case loss (in terms of social utility), this does not mean that the two algorithms are equivalent: we show in particular that the sequential algorithm has several drawbacks, among which the lack of swap stability is certainly the most annoying, as agents can rather easily notice it, and (justifiably) ask for explanations. An easy patch is thus to let agents swap until a stable allocation is reached – but is there such a guarantee? We show that this is the case, despite the fact that swaps may actually decrease social welfare. In other words, stability comes at a price. We provided an upper bound on this price that supports the stability requirement since the worst-case price is constant and quite low. As suggested in the introduction, another compelling argument in favour of a stable mechanism is indeed that in the absence of this property *agents may develop their own system to perform swaps*. However, such a system would be unlikely to be inclusive, raising concerns of fairness – did anyone really get a chance to swap? This issue happened in practice in the Netherland in 2015, as reported in [21], when “an Amsterdam court ruled (...) that students going to high school aren’t allowed to trade places with each other at different schools.” Our results suggest that it may be possible to avoid this issue by allowing swaps in a second phase after the sequential mechanism, as long as they are implemented through a legitimate platform guaranteeing equal access to all users. A challenging issue is to identify actual swap opportunities, but this can be done by eliciting preferences from users on a voluntary basis. In France, web platforms have recently been developed to make easier the exchange of lands between farmers[4] and the exchange of public housings in Ile de France[5]. These websites connect the agents and make swaps easier. A related issue which we didn't touch upon here is the problem of manipulability. Finally, there are of course several other metrics which could refine our understanding of the relative merits of these algorithms, for instance it would be good to study various fairness measures. Another interesting issue would be to investigate settings where different agents may have different homophily biases, i.e. different utility trade-offs φ. In Appendix F we show how such problems can be mapped to the setting where φ is the same for all agents, and discuss which results gets affected by this translation. Finally, while diversity is only assessed through the satisfaction of quotas in this paper, a more ambitious research agenda would be to investigate the interplay between (a metric of) diversity and utility, for instance by adjusting the quota thresholds.

REFERENCES

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

[2] Aishwarya Agarwal, Edith Elkind, Jiarui Gan, and Alexandros A. Voudouris. Swap stability in schelling games on graphs. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence, AAAI 2020*, pages 1758–1765, 2020.

[3] Orhan Aygun and Inácio Bó. College admission with multidimensional privileges: The brazilian affirmative action case. *Game Theory & Bargaining Theory eJournal*, 2019.

[4] Haris Aziz, Florian Brandl, Felix Brandt, Paul Harrenstein, Martin Olsen, and Dominik Peters. Fractional hedonic games. *ACM Trans. Econ. Comput.*, 7(2), June 2019.

[5] Haris Aziz. Developments in multi-agent fair allocation. In *Proceedings of The 34th AAAI Conference on Artificial Intelligence, AAAI 2020*, pages 13563–13568, New York, NY, USA, 2020. AAAI Press.

[6] Nawal Benabbou, Mithun Chakraborty, Xuan-Vinh Ho, Jakub Sliwinski, and Yair Zick. Diversity constraints in public housing allocation. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS'18, page 973–981, 2018.

[7] Anna Bogomolnaia and Hervé Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295 – 328, 2001.

[8] Sylvain Bouveret, Katarína Cechlárová, Edith Elkind, Ayumi Igarashi, and Dominik Peters. Fair division of a graph. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI'17*, pages 135–141, 2017.

[9] Robert Bredereck, Edith Elkind, and Ayumi Igarashi. Hedonic diversity games. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS'19, page 565–573, 2019.

[10] Inácio Bó. Fair implementation of diversity in school choice. *Games and Economic Behavior*, 97:54 – 63, 2016.

[11] Ankit Chauhan, Pascal Lenzner, and Louise Molitor. Schelling segregation with strategic agents. In *Algorithmic Game Theory*, pages 137–149. Springer International Publishing, 2018.

[12] Anastasia Damamme, Aurélie Beynier, Yann Chevaleyre, and Nicolas Maudet. The power of swap deals in distributed resource allocation. In *Proceedings of 14th International Conference on Au-tonomous Agents and MultiAgent System*, 2015.

[13] Edith Elkind, Jiarui Gan, Ayumi Igarashi, Warut Suksompong, and Alexandros A. Voudouris. Schelling games on graphs. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI'19*, pages 266–272, 7 2019.

[14] Edith Elkind, Neel Patel, Alan Tsang, and Yair Zick. Keeping your friends close: Land allocation with friends. In Christian Bessiere, editor, *Proceedings of the 29th International Joint Conference on Artificial Intelligence, IJCAI'20*, pages 318–324, 7 2020.

[15] Ulle Endriss, Nicolas Maudet, Fariba Sadri, and Francesca Toni. Negotiating socially optimal allocations of resources. *Journal of artificial intelligence research*, 2006.

[16] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In Christoph Meinel and Sophie Tison, editors, *Proceedings of the 16th Symposium on Theoretical Aspects of Computer Science, STACS*, pages 404–413, 1999.

[17] Bo Li and Yingkai Li. Fair resource sharing and dorm assignment. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, page 708–716, Richland, SC, 2020. International Foundation for Autonomous Agents and Multiagent Systems.

[18] Jiamou Liu and Ziheng Wei. Network, popularity and social cohesion: A game-theoretic approach. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence, AAAI'17*, pages 600–606, 2017.

[19] Sagar Massand and Sunil Simon. Graphical one-sided markets. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI'19*, 2019.

[4] https://www.echangeparcelle.fr

[5] https://www.echangerhabiter.fr

[20] M. Olsen. On defining and computing communities. In *Proceedings of the 18th Computing: Australian Theory Symposium (CATS)*, pages 97–102, 2012.

[21] Ronald de Haan. Why a Dutch court stopped high school students from swapping schools. https://medium.com/social-choice/why-a-dutch-court-stopped-high-school-students-from-\exchanging-schools-1315303a48b6, 2017. Accessed: 2020-10-08.

[22] Alvin E. Roth. The economist as engineer: Game theory, experimental economics and computation as tools of design economics. *Econometrica*, 4:1341–1378, July 2002.

[23] Thomas C. Schelling. Models of segregation. *The American Economic Review*, 59(2):488–493, 1969.

[24] Toby Walsh. Fair division: The computer scientist's perspective. In Christian Bessiere, editor, *Proceedings of the 29th International Joint Conference on Artificial Intelligence, IJCAI 2020*, pages 4966–4972, 2020.

[25] Xiaoming Zheng and Sven Koenig. K-swaps: Cooperative negotiation for solving task-allocation problems. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI'09*, pages 373–379, 2009.

[26] Kolos Csaba Ágoston, Péter Biró, and Richárd Szántó. Stable project allocation under distributional constraints. *Operations Research Perspectives*, 5(C):59–68, 2018.

A MIXED-INTEGER-LINEAR PROGRAMMING FORMULATIONS

In order to compute a valid allocation of maximum utilitarian social welfare, one can use the following mixed-integer linear program (involving only a polynomial number of variables/constraints):

$$
\begin{aligned}
\max \quad & \sum_{i \in \mathcal{N}} \sum_{h \in \mathcal{M}} x_{ih} \times u_i^I(h) + \varphi \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{T}(i)} \sum_{q \in [l]} \frac{s_{ijq}}{|B_q|} & & (1) \\
\text{s. t.} \quad & \sum_{h \in \mathcal{M}} x_{ih} \leq 1 & \forall i \in \mathcal{N} \quad & (2) \\
& \sum_{i \in \mathcal{N}} x_{ih} = 1 & \forall h \in \mathcal{M} \quad & (3) \\
& \sum_{i \in T_p} \sum_{h \in B_q} x_{ih} \leq \lambda_{p,q} & \forall p \in [k],\ q \in [l] \quad & (4) \\
& b_{iq} = \sum_{h \in B_q} x_{ih} & \forall i \in \mathcal{N},\ q \in [l] \quad & (5) \\
& s_{ijq} \leq b_{iq} & \forall i, j \in \mathcal{N},\ q \in [l] \quad & (6) \\
& s_{ijq} \leq b_{jq} & \forall i, j \in \mathcal{N},\ q \in [l] \quad & (7) \\
& x_{ih} \in \{0,1\}, b_{iq} \in \{0,1\}, s_{ij,} \in \{0,1\} & \forall i, j \in \mathcal{N},\ h \in \mathcal{M} \quad & (8)
\end{aligned}
$$

where x_{ih}, $b_{i,q}$ and s_{ijq}, with $i, j \in \mathcal{N}$, $q \in [l]$, are the variables of the program (see Equation (8)). Variable x_{ih} is equal to one if item h is assigned to agent i. Variable $b_{i,q}$ is equal to 1 if agent i is assigned to an item in block B_q. Variable s_{ijq} is equal to one if agents i and j are both assigned to block B_q. Equation (2) is used to ensure that every agent is assigned to at most one item, while Equation (3) ensures that every item is assigned to one agent. Equation (4) corresponds to type-block quotas. Equation (5) imposes that variables b_{iq} are equal to 1 if and only if agent i is assigned to an item in block B_q. Equations (6-7) are used to ensure that boolean variables s_{ijq} are equal to 0 when agents i or j are not in B_q, and to 1 otherwise (since the objective function is to be maximized).

Now we propose the following mixed-integer program for finding a valid stable allocation of maximum utilitarian social welfare.

$$\max \sum_{i \in \mathcal{N}} \sum_{h \in \mathcal{M}} x_{ih} \times u_i^I(h) + \varphi \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{T}(i)} \sum_{q \in [l]} \frac{s_{ijq}}{|B_q|} \tag{9}$$

$$\text{s.t.} \sum_{h \in \mathcal{M}} x_{ih} = 1 \qquad \forall i \in \mathcal{N} \tag{10}$$

$$\sum_{i \in \mathcal{N}} x_{ih} = 1 \qquad \forall h \in \mathcal{M} \tag{11}$$

$$\sum_{i \in T_p} \sum_{h \in B_q} x_{ih} \leq \lambda_{pq} \qquad \forall p \in [k],\ q \in [l] \tag{12}$$

$$b_{iq} = \sum_{h \in B_q} x_{ih} \qquad \forall i \in \mathcal{N},\ q \in [l] \tag{13}$$

$$s_{ijq} \leq b_{iq} \qquad \forall i, j \in \mathcal{N},\ q \in [l] \tag{14}$$

$$s_{ijq} \leq b_{jq} \qquad \forall i, j \in \mathcal{N},\ q \in [l] \tag{15}$$

$$c_{ij1} \geq \sum_{h \in \mathcal{M}} (x_{jh} - x_{ih}) \times u_i^I(h) + \varphi \sum_{q \in [l]} \left(\frac{b_{jq} - s_{ijq}}{|B_q|} + \sum_{a \in \mathcal{T}(i) \setminus \{j\}} \frac{s_{jaq}}{|B_q|} - \sum_{a \in \mathcal{T}(i)} \frac{s_{iaq}}{|B_q|} \right) \qquad \forall i, j \in \mathcal{N} \tag{16}$$

$$c_{ij2} \geq \sum_{h \in \mathcal{M}} (x_{ih} - x_{jh}) \times u_j^I(h) + \varphi \sum_{q \in [l]} \left(\frac{b_{iq} - s_{ijq}}{|B_q|} + \sum_{a \in \mathcal{T}(j) \setminus \{i\}} \frac{s_{iaq}}{|B_q|} - \sum_{a \in \mathcal{T}(j)} \frac{s_{jaq}}{|B_q|} \right) \qquad \forall i, j \in \mathcal{N} \tag{17}$$

$$c_{ij3} \times \Lambda \geq \sum_{q \in [l]} \left(\lambda_{T(i)q} \times b_{jq} - \sum_{a \in T(i)} s_{jaq} \right) \qquad \forall i, j \in \mathcal{N}, j \notin \mathcal{T}(i) \tag{18}$$

$$c_{ij4} \times \Lambda \geq \sum_{q \in [l]} \left(\lambda_{T(j)q} \times b_{iq} - \sum_{a \in T(j)} s_{iaq} \right) \qquad \forall i, j \in \mathcal{N}, j \notin \mathcal{T}(i) \tag{19}$$

$$\sum_{r=1}^{4} c_{ijr} \leq 3 \qquad \forall i, j \in \mathcal{N} \tag{20}$$

$$x_{ih} \in \{0, 1\} \qquad \forall i \in \mathcal{N},\ h \in \mathcal{M} \tag{21}$$

$$b_{iq} \in \{0, 1\}, s_{ijq} \in \{0, 1\} \qquad \forall i, j \in \mathcal{N},\ q \in [l] \tag{22}$$

$$c_{ijr} \in \{0, 1\} \qquad \forall i, j \in \mathcal{N},\ r \in [4] \tag{23}$$

where x_{ih}, b_{iq}, s_{ijq}, and c_{ijr}, with $i, j \in \mathcal{N}$, $q \in [l]$, $r \in [4]$, are the variables of the program, and Λ is a constant defined by $\Lambda = \max_{p \in [k], q \in [l]} \lambda_{pq}$. Variables x_{ih}, b_{iq} and s_{ijq} are defined as before. Boolean variables c_{ijr}, with $r \in [4]$, are used to ensure stability. More precisely, for any two agents $i, j \in \mathcal{N}$:

- c_{ij1} is equal to 1 if and only if i prefers the item assigned to agent j to hers (Equation (16)).
- c_{ij2} is equal to 1 if and only if j prefers the item assigned to agent i to hers (Equation (17)).

- c_{ij3} is equal to 1 if and only if type-block quotas allow agent i to receive the item assigned to agent j (Equation (18)).
- c_{ij4} is equal to 1 if and only if type-block quotas allow agent j to receive the item assigned to agent i (Equation (19)).

Equation (20) is used to check that at least one condition is not verified. If they are all verified, then there is a valid improving swap deal between i and j, which means that the allocation is not stable.

B PROOF OF LEMMA 1

Proof. First, let us prove that the largest utility ratio between any two valid allocations is unbounded in the general case. Given a parameter $k \in \mathbb{N}$, we consider a family of problem instances $\mathcal{I}_k$ with k types of agents, k agents per type (i.e. $|T_p| = k$ for all $p \in [k]$), k blocks of items, k items per block (i.e. $|B_q| = k$ for all $q \in [l]$) and type-block capacities $\lambda_{p,q} = k$ for all $p \in [k]$ and $q \in [l]$. Thus, we have $n = k^2$ agents and $m = k^2$ items in total, and type-block quotas are not restrictive.

Let each type of agents and each block of goods be ordered so that we can call $a_{p,i}$ the agent number i of type p and $b_{q,j}$ the good number j of block q.

Assume that the agents' item-based utilities are defined by:

$$u^I_{p,i}(A) = \begin{cases} 1 & \text{if item } A(p,i) = b_{q,j} \text{ is in such that } p = q \text{ and } i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

and $\varphi \neq 0$. Obviously, any allocation A such that $A(p,i) = b_{p,i+1[k]}$ for all agents $i \in \mathcal{N}$ is valid and maximizes the utilitarian social welfare (every agent receives an item with top utility and blocks are only composed of agents of the same type). Let $A^*_{\mathcal{I}_k}$ be some optimal allocation of instance $\mathcal{I}_k$. By definition, we have:

$$\begin{aligned} sw(A^*_{\mathcal{I}_k}) &= \sum_{p=1}^{k}\sum_{i=1}^{k}\left(u^I_{p,i}(A^\top_{\mathcal{I}_k}) + \varphi u^N_{p,i}(A^\top_{\mathcal{I}_k})\right) \\ &= \sum_{i=1}^{k^2}(1 + \varphi \times 1) \\ &= \sum_{i=1}^{k^2}(1 + \varphi) \\ &= k^2(1 + \varphi) \end{aligned}$$

Now consider any allocation A_k such that each block only includes one agent per type, and such that, in each bloc q, the agent $a_{q,i}$ is allocated to a good $b_{q,j}$ with $j = i$. Allocation A_k is obviously valid, and we have:

$$\begin{aligned} sw(A_k) &= \sum_{p=1}^{k}\sum_{i=1}^{k}\left(u^I_{p,i}(A_k) + \varphi u^N_{p,i}(A_k)\right) \\ &= \sum_{p=1}^{k}\sum_{i=1}^{k}\left(0 + \varphi \times \frac{1}{\mathcal{B}(A_k(p,i))}\right) \\ &= \sum_{i=1}^{k^2}\varphi\frac{1}{k} \\ &= \varphi k \end{aligned}$$

Therefore, for instance $\mathcal{I}_k$, we have:

$$\frac{sw(A^*_{\mathcal{I}_k})}{sw(A_k)} = \frac{k^2(1+\varphi)}{\varphi k} = \frac{1+\varphi}{\varphi}k$$

Since this utility ratio grows linearly with k, the largest utility ration between any two valid allocations is unbounded in the general case.

Now, let us prove that $\frac{1+\varphi}{\varphi}k$ is a tight upper bound when k is a constant and $\varphi > 0$. To do so, it is now sufficient to prove that the utility ratio between any two valid allocations cannot be greater than this value. Note that the value $sw_{up} := (1+\varphi) \times m$ is an obvious upper bound on the utility of a valid allocation (each item is assigned to an agent with an item-based utility equal to 1 and blocks only include agents of the same type). We now focus on the determination of a lower bound on the utility of a valid allocation. Since the item-based utility value can be equal to zero for all agents simultaneously (when they all receive an item for which they have no utility for it), we only need to find a lower bound on the sum of the agents' neighbour-utilities. Let us focus on a block B_q, with $q \in [l]$. For any valid allocation A, the sum of the neighbour-utilities in block B_q is given by:

$$\begin{aligned} \sum_{\substack{i\in\mathcal{N}\\ A(i)\in B_q}} u^N_i(A) &= \sum_{\substack{i\in\mathcal{N}\\ A(i)\in B_q}}\ \sum_{\substack{j\in\mathcal{N}\\ A(j)\in B_q}} \frac{\mathbb{I}(\mathcal{T}(i),\mathcal{T}(j))}{|B_q|} \\ &= \frac{1}{|B_q|}\sum_{p=1}^{k}|\{i \in T_p : A(i) \in B_q\}|^2 \end{aligned}$$

Therefore, we can obtain a lower bound on this value by solving the following convex optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{|B_q|}\sum_{p=1}^{k} x^2_{p,q} \\ s.t. \quad & \sum_{p=1}^{k} x_{p,q} = |B_q| \\ & x_{p,q} \in \mathbb{R}_+, \forall p \in [k] \end{aligned}$$

where $x_{p,q}$ is a continuous variable representing the number of agents of type T_p who are assigned to an item in block B_q, for all $p \in [k]$. The corresponding optimal solution x^* is such that:

$$\exists\lambda \in \mathbb{R},\ \nabla f(x^*) + \lambda\nabla g(x^*) = 0$$

where $f(x)=\frac{1}{|B_q|}\sum_{p=1}^{k} x_{p,q}^2$ and $g(x)=\left(\sum_{p=1}^{k} x_{p,q}\right) - |B_q|$. Thus we know that x^* and λ are such that:

$$\frac{2}{|B_q|}\begin{pmatrix} x_1^* \\ \vdots \\ x_k^* \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

and therefore $x_p^* = -\frac{\lambda|B_q|}{2}$ for all $p \in [k]$. Since $\sum_{p=1}^{k} x_k^* = |B_q|$ also holds, we obtain $\lambda = -\frac{2}{k}$ and then $x_p^* = \frac{|B_q|}{k}$ for all $p \in [k]$. We finally obtain the following lower bound on the sum of the neighbour-utilities in block B_q:

$$\sum_{\substack{i\in\mathcal{N}\\ A(i)\in B_q}} u_i^N(A) \geq f(x^*) = \frac{1}{|B_q|}\sum_{p=1}^{k}\left(\frac{|B_q|}{k}\right)^2 = \frac{|B_q|}{k}$$

for any valid allocation A. Summing over all blocks B_q, $q \in [l]$, we obtain a lower bound on the sum of the neighbour-utilities:

$$\sum_{i\in\mathcal{N}} u_i^N(A) = \sum_{q=1}^{l} \sum_{\substack{i\in N\\ A(i)\in B_q}} u_i^N(A) \geq \sum_{q=1}^{l} \frac{|B_q|}{k} = \frac{1}{k}\sum_{q=1}^{l} |B_q| = \frac{m}{k}$$

Thus for any valid allocation A, we have:

$$sw(A) = \sum_{i\in\mathcal{N}} u_i(A) = \sum_{i\in\mathcal{N}} \left(u_i^I(A) + \varphi u_i^N(A)\right) \geq 0 + \varphi\frac{m}{k}$$

and therefore $sw_{low} := \varphi\frac{m}{k}$ is a lower bound on the utility of any valid allocation. Hence the utility ratio between any two valid allocations cannot be greater than $\frac{sw_{up}}{sw_{low}} = \frac{1+\varphi}{\varphi}k$. □

C DETAILED PROOF OF PROPOSITION 1

Proof. Here we only need to prove that any valid allocation that maximizes the utilitarian social welfare is stable. For the case of item-focused agents, the argument is straightforward: if two agents wish to swap and they are allowed to do so, then implementing the swap would strictly augment the utilities of both agents, while leaving the other utilities unchanged, which yields an allocation with a strictly larger utility. Therefore any valid allocation maximizing the social welfare is necessarily stable.

For neighbour-focused agents, we now prove that if two agents wish to swap to increase their individual utilities, the utilitarian social welfare will increase by twice as much. For a given non-stable allocation A, consider a valid improving swap-deal between some agent $i_1 \in T_{p_1}$ with $A(i_1) \in B_{q_1}$ and some agent $i_2 \in T_{q_2}$ with $A(i_2) \in B_{q_2}$. Since the item-based component of their utilities is equal to zero, we know that these two agents are of different types and belong to different blocks (i.e., we necessarily have $p_1 \neq p_2$ and $q_1 \neq q_2$). Let ΔU be the utility difference between A' the allocation resulting from the swap-deal and allocation A, i.e. $\Delta U = sw(A') - sw(A)$. Let us decompose ΔU as follows:

$$\Delta U = \Delta U_1 + \Delta U_2$$

where ΔU_1 (resp. ΔU_2) is the change of utility for the agents that are involved (resp. not involved) in the swap-deal. Since the composition of blocks B_{q_1} and B_{q_2} changes, the swap may incur a loss or a gain in utility on the agents in both blocks. This externalities are measured by U_2. For all $p \in [k]$ and $q \in [l]$, let $A_{p,q}$ denote the number of agents of type T_p in block B_q. Using these notations, the swap-deal modifies the utilities of the two swapping agents as follows:

$$
\begin{aligned}
\Delta U_1 &= u_{i_1}(A') - u_{i_1}(A) + u_{i_2}(A') - u_{i_2}(A) \\
&= \varphi\Big(\frac{A_{p_1,q_2}+1}{|B_{q_2}|} - \frac{A_{p_1,q_1}}{|B_{q_1}|} + \frac{A_{p_2,q_1}+1}{|B_{q_1}|} - \frac{A_{p_2,q_2}}{|B_{q_2}|}\Big) \\
&= \varphi\Big(\frac{A_{p_2,q_1} - A_{p_1,q_1} + 1}{|B_{q_1}|} + \frac{A_{p_1,q_2} - A_{p_2,q_2} + 1}{|B_{q_2}|}\Big)
\end{aligned}
$$

We now prove that $\Delta U_2 = \Delta U_1$:

$$
\begin{aligned}
\Delta U_2 = &\sum_{\substack{i \in T_{p_1}\setminus\{i_1\} \\ A(i)\in B_{q_1}\cup B_{q_2}}} \Big(u_i(A') - u_i(A)\Big) + \sum_{\substack{i \in T_{p_2}\setminus\{i_2\} \\ A(i)\in B_{q_1}\cup B_{q_2}}} \Big(u_i(A') - u_i(A)\Big) \\
&= (A_{p_1,q_1} - 1) \times \Big(\frac{A_{p_1,q_1}-1}{|B_{q_1}|} - \frac{A_{p_1,q_1}}{|B_{q_1}|}\Big)\varphi + A_{p_1,q_2}\Big(\frac{A_{p_1,q_2}+1}{|B_{q_2}|} - \frac{A_{p_1,q_2}}{|B_{q_2}|}\Big)\varphi \\
&+ A_{p_2,q_1}\Big(\frac{A_{p_2,q_1}+1}{|B_{q_1}|} - \frac{A_{p_2,q_1}}{|B_{q_1}|}\Big)\varphi + (A_{p_2,q_2} - 1)\Big(\frac{A_{p_2,q_2}-1}{|B_{q_2}|} - \frac{A_{p_2,q_2}}{|B_{q_2}|}\Big)\varphi \\
&= (A_{p_1,q_1} - 1)\frac{-1}{M_{q_1}}\varphi + A_{p_1,q_2}\frac{1}{M_{q_2}}\varphi + A_{p_2,q_1}\frac{1}{M_{q_1}}\varphi + (A_{p_2,q_2} - 1)\frac{-1}{M_{q_2}}\varphi \\
&= u_{i_1}(A') - u_{i_1}(A) + u_{i_2}(A') - u_{i_2}(A) \\
&= \Delta U_1
\end{aligned}
$$

Thus $\Delta U = 2\Delta U_1$, which means that any improving swap-deal would strictly increase the social welfare. This implies that any valid allocation maximizing the social welfare is necessarily stable. □

D PROOF OF PROPOSITION 2

PROOF. We need to show that we have $PoS \leq 2$ and $PoS > 1.6$.

$PoS \leq 2$. From Example 1, we know that there exist instances for which no optimal allocation is stable. Let us consider such an instance, and let OPT denote such an optimal allocation. From this allocation, we can apply the SWAP-mechanism (presented in Section 5) which consists in performing bilateral improving swap-deals until reaching a stable outcome (see Proposition 7). Let $POST$ be such an equilibrium. Our aim is to prove that the inequality $\frac{sw(OPT)}{sw(POST)} \leq 2$ holds. To do so, let us decompose $sw(OPT)$ and $sw(POST)$ as follows:

$$
\begin{aligned}
sw(OPT) &= U^I_{OPT} + \varphi U^N_{OPT} \\
sw(POST) &= U^I_{POST} + \varphi U^N_{POST}
\end{aligned}
$$

where U^I_{OPT} (resp. U^I_{POST}) is the sum of the item-based utilities of the agents in the allocation OPT (resp. $POST$), and U^N_{OPT} (resp. U^N_{POST}) is the sum of their neighbour-based utilities in the same allocation. We now use pot the potential function introduced in the proof of Proposition 7, which is defined by:

$$
pot(A) = sw(A) + \sum_{i \in \mathcal{N}} u^I_i(A)
$$

In particular, we have $pot(OPT) = sw(OPT) + U^I_{OPT}$ and $pot(POST) = sw(POST) + U^I_{POST}$. In the proof of Proposition 7, it is shown that pot strictly increases when performing any valid improving swap. Since $POST$ is obtained from

OPT by performing at least one valid improving swap, we have $pot(POST) > pot(OPT)$, i.e. $sw(POST) + U^I_{POST} > sw(OPT) + U^I_{OPT}$.

Moreover, we have $sw(OPT) > sw(POST)$ since OPT is optimal (and no optimal allocation is stable by hypothesis). Therefore, we necessarily have $U^I_{OPT} < U^I_{POST}$. Now, note that we have:

$$\frac{sw(OPT)}{sw(POST)} = \frac{sw(OPT) + sw(POST) - sw(POST)}{sw(POST)} = 1 + \frac{sw(OPT) - sw(POST)}{sw(POST)}$$

Moreover, we have:

$$\begin{aligned} & sw(OPT) - sw(POST) \\ & \quad = sw(OPT) - sw(POST) + U^I_{OPT} - U^I_{OPT} \\ & \quad = pot(OPT) - sw(POST) - U^I_{OPT} + U^I_{POST} - U^I_{POST} \\ & \quad = pot(OPT) - pot(POST) + U^I_{POST} - U^I_{OPT} \end{aligned}$$

Therefore, we obtain:

$$\begin{aligned} \frac{sw(OPT)}{sw(POST)} &= 1 + \frac{pot(OPT) - pot(POST) + U^I_{POST} - U^I_{OPT}}{sw(POST)} \\ &< 1 + \frac{U^I_{POST} - U^I_{OPT}}{sw(POST)} \quad (\text{since } pot(OPT) < pot(POST)) \\ &< 1 + \frac{U^I_{POST}}{sw(POST)} \quad (\text{since } U^I_{OPT} \geq 0) \\ &< 2 \quad (\text{since } sw(POST) = U^I_{POST} + \varphi U^N_{POST}) \end{aligned}$$

This means that any equilibrium that is reachable from the optimal allocation is at least half as good as the optimum. Since the best equilibrium is at least as good as any reachable equilibrium, it is also at least half as good as the optimum. Thus the utility ratio between any optimal allocation and any equilibrium maximizing the social welfare is strictly smaller than 2 for any instance of our allocation problem with diversity constraints. Since PoS is the least upper bound on this utility ratio over all possible instances, we can conclude that $PoS \leq 2$ holds.

$PoS > 1.6$. We show here that the price of stability is at least 1.6, by providing an instance for which the utility ratio between the optimal and the best (valid) stable allocations reaches this value. More precisely, we consider an instance with 16 agents partitioned into 4 types: $T_1 = \{a_{1,1}, a_{1,2}, a_{1,3}, a_{1,4}\}$, $T_2 = \{a_{2,1}, a_{2,2}, a_{2,3}, a_{2,4}\}$, $T_3 = \{a_{3,1}, a_{3,2}, a_{3,3}, a_{3,4}\}$, and $T_4 = \{a_{4,1}, a_{4,2}, a_{4,3}, a_{4,4}\}$. It involves 16 items partitioned into 4 blocks: $B_1 = \{h_{1,1}, h_{1,2}, h_{1,3}, h_{1,4}\}$, $B_2 = \{h_{2,1}, h_{2,2}, h_{2,3}, h_{2,4}\}$, $B_3 = \{h_{3,1}, h_{3,2}, h_{3,3}, h_{3,4}\}$, and $B_4 = \{h_{4,1}, h_{4,2}, h_{4,3}, h_{4,4}\}$. Type-block quotas are defined by $\lambda_{pq} = 4$ when $p = q$, and $\lambda_{pq} = 1$ otherwise. The utility trade-off φ is equal to 1. The item-based utilities are defined by the matrix P_ϵ, with $\epsilon > 0$, as follows:

$$\forall i \in T_p, h \in B_q,\ u^I_i(h) = P_{p,q} \text{ where } P_\epsilon = \begin{pmatrix} 0 & \frac{3}{4}+\epsilon & \frac{2}{4}+\epsilon & \frac{1}{4}+\epsilon \\ \frac{3}{4}+\epsilon & 0 & \frac{1}{4}+\epsilon & \frac{2}{4}+\epsilon \\ \frac{2}{4}+\epsilon & \frac{1}{4}+\epsilon & 0 & \frac{3}{4}+\epsilon \\ \frac{1}{4}+\epsilon & \frac{2}{4}+\epsilon & \frac{3}{4}+\epsilon & 0 \end{pmatrix}$$

It can be shown[6] that:

- the social welfare is maximized by any allocation assigning all the agents of type T_p in block B_p, for $p \in \{1, \dots, 4\}$. In such an allocation, the utility of each agent is $0 + \varphi \times \frac{4}{4} = 1$ and the social welfare is equal to 16. Note that such an allocation is not stable, e.g., agents of type T_1 and agents of type T_2 want to swap with their items to increase their utilities from 1 to $1 + \varepsilon$
- any allocation assigning one agent of each type in every block is a valid stable allocation of maximum social welfare. Its social welfare is equal to $4 \times (0 + \varphi \times \frac{1}{4}) + 4 \times (\frac{1}{4} + \varepsilon + \varphi \times \frac{1}{4}) + 4 \times (\frac{2}{4} + \varepsilon + \varphi \times \frac{1}{4}) + 4 \times (\frac{3}{4} + \varepsilon + \varphi \times \frac{1}{4}) = 10 + 12\varepsilon$.

Hence, the utility ratio between any optimal allocation and any valid stable allocation of maximum social welfare is equal to $\frac{16}{10+12\varepsilon}$, which tends towards 1.6 when ε tends towards 0.

□

E PROOF OF PROPOSITION 8

PROOF. Given a parameter $x \in \mathbb{N}$, we consider a family of instances with k types of agents, each type containing $xk - 1$ agents: for all $p \in [k]$, $T_p = \{a_{p,1}, a_{p,2}, \dots, a_{p,xk-1}\}$. Similarly, we consider k blocks, each of them including $xk - 1$ items: for all $q \in [k]$, $B_q = \{h_{q,1}, h_{q,2}, \dots, h_{q,xk-1}\}$. For all $p, q \in [k]$, the type-quota capacities are defined by:

$$\lambda_{p,q} = \begin{cases} xk - 1 & \text{if } p = q \\ x & \text{otherwise} \end{cases}$$

Now, for any allocation A, we define the agents' utilities as follows: for all types $p \in [k]$ and all $a_{p,i} \in T_p$, the item-based utility of agent $a_{p,i}$ is given by:

$$u^I_{a_{p,i}}(A) = \begin{cases} 1 & \text{if } A(a_{p,i}) = h_{p,i} \\ 0 & \text{otherwise} \end{cases}$$

Here the optimal allocation A_{best} is simply defined by $A_{best}(a_{p,i}) = h_{p,i}$ for all $p \in [k]$ and all $a_{p,i} \in T_p$, which consists in grouping the agents according to their types, while assigning every item to the single agent with a utility of 1 for it. The utilitarian social welfare of A_{best} is obviously:

$$\begin{aligned} sw(A_{best}) &= \sum_{p=1}^{k} \sum_{i=1}^{xk-1} \left(u^I_{a_{p,i}}(A) + \varphi u^N_{a_{p,i}}(A)\right) \\ &= \sum_{p=1}^{k} \sum_{i=1}^{xk-1} (1 + \varphi) \\ &= k(xk - 1)(1 + \varphi) \end{aligned}$$

We now build an allocation, called A_{worst}, that minimizes the utilitarian social over all possible *stable* valid allocations. For this, we first rename the agent: each agent of type T_p is identified by a triplet $(p, q, y) \in [k]^2 \times [x] \setminus \{(z, z, x) : \forall z \in [k]\}$. Similarly, each item of block B_p is identified by a triplet $(p, q, y) \in [k]^2 \times [x] \setminus \{(z, z, x) : \forall z \in [k]\}$. By doing so, every item gives a utility of 1 to the agent who shares the same triplet. Now, for all $p \in [k]$ and for all $y \in [x - 1]$, we set:

$$A_{worst}(a_{(p,p,y)}) = \begin{cases} h_{(p,p,y-1)} & \text{if } y \in \{2, \dots, x - 1\}, \\ h_{(p,p,x-1)} & \text{otherwise (i.e. when } y = 1). \end{cases}$$

[6] These results can be shown using our mixed-integer linear programming formulations given in Appendix A, or by enumerating all the valid allocations.

Thus $x-1$ agents of each type T_p are assigned to $x-1$ items of block B_p, but none of them receives her favourite item. Moreover, there is no pair of agents such that both agents would be happy to swap due to the cycling structure.

To allocate the remaining agents (i.e. agents $a_{(p,q,y)}$ with $p, q \in [k]$, $p \neq q$, and $y \in [x]$), we consider k one-to-one correspondences σ_p, with $p \in [k]$, such that $\sigma_p : \{1, \dots, k\} \backslash \{p\} \to \{1, \dots, k\} \backslash \{p\}$ and $\sigma_p(q) \neq q$ for all $q \in [k] \backslash \{p\}$. Then, for all $p, q \in [k]$, with $q \neq p$, and for all $y \in [x]$, we set:

$$A_{worst}(a_{(p,q,y)}) = h_{(q,\sigma_p(q),y)}$$

Thus, for each type T_p with $p \in [k]$, the remaining $(k-1)x$ agents are distributed evenly among the blocks B_q, with $q \neq p$, so that each of them includes exactly x agents of type T_p. It is easy to check that A_{worst} satisfies type-block quotas, which implies that it is a valid allocation. Moreover, for each type $p \in [k]$ and each block B_q, with $q \neq p$, the corresponding x agents would be happy to swap their items only with the agents in block B_p that own their favourite items, but type-block quotas prevent such swaps from happening (we know that they are not of types p or q by definition of σ_p). Therefore, A_{worst} is a valid stable allocation, whose social welfare is:

$$\begin{aligned}
sw(A_{worst}) &= \sum_{p=1}^{k} \sum_{i=1}^{xk-1} \left(u^I_{a_{p,i}}(A_{worst}) + \varphi u^N_{a_{p,i}}(A_{worst}) \right) \\
&= \sum_{p=1}^{k} \sum_{i=1}^{xk-1} \left(0 + \varphi u^N_{a_{p,i}}(A_{worst}) \right) \\
&= \varphi \sum_{p=1}^{k} \left(\sum_{i=1}^{x-1} u^N_{a_{p,i}}(A_{worst}) + \sum_{i=x}^{kx-1} u^N_{a_{p,i}}(A_{worst}) \right) \\
&= \varphi \sum_{p=1}^{k} \left(\sum_{i=1}^{x-1} \frac{x-1}{kx-1} + \sum_{i=x}^{kx-1} \frac{x}{kx-1} \right) \\
&= \varphi \sum_{p=1}^{k} \left(\frac{(x-1)^2}{xk-1} + \frac{x^2(k-1)}{xk-1} \right) \\
&= \frac{\varphi k (x^2 k - 2x + 1)}{xk-1}
\end{aligned}$$

It follows that the largest utility ratio between any two valid stable allocations is:

$$\frac{sw(A_{best})}{sw(A_{worst})} = \frac{(1+\varphi)(xk-1)k}{\frac{\varphi k (x^2 k - 2x + 1)}{xk-1}} = \frac{(1+\varphi)(xk-1)^2}{\varphi (x^2 k - 2x + 1)}$$

When x tends towards $+\infty$, this ratio tends towards $\frac{(1+\varphi)(xk)^2}{\varphi x^2 k} = \frac{(1+\varphi)k}{\varphi}$, which implies that the bound is reached at the limit. □

F NEIGHBOURHOOD-BASED UTILITIES WITH DIFFERENT UTILITY TRADE-OFFS

In the main part of the paper, we assumed that all the agents have the same utility trade-off φ balancing the relative importance of the item-based utility and the neighbour-based utility. In this section, we discuss whether instances where the agents have different utility trade-offs can be translated to a single utility trade-off instance. More precisely, here we consider instances where each agent i has its own utility trade-off $\varphi_i \in (0, 1]$, and item-based utilities are normalized, i.e. each agent i has at least one item $h \in \mathcal{M}$ such that $u^I_i(h) = 0$ and at least one item $h \in \mathcal{M}$ such that $u^I_i(h) = 1$. We first show how to convert such an instance to another one with a unique utility trade-off. We then discuss the convergence of the swap mechanism and investigate the validity of our worst-case results.

Instance transformation. From an instance $\mathcal{I}_1$ where the agents have different utility trade-offs, we define a new instance $\mathcal{I}_2$ with the same components $\mathcal{N}, \mathcal{M}, B, T, \lambda$ where the agents have the same utility trade-off $\varphi = min_{i \in \mathcal{N}} \varphi_i$. More precisely, we consider new utilities $(v_i)_{i \in \mathcal{N}}$ defined by:

$$v_i(A) = v_i^I(A) + \varphi v_i^N(A)$$

for every allocation A, where $v_i^I(A) = \frac{\varphi}{\varphi_i} u_i^I(A)$ and $v_i^N(A) = u_i^N(A)$.

On the convergence of the swap-deal mechanism. Note that $v_i(A)$ is simply obtained by multiplying $u_i(A)$ by a positive value. Therefore, u_i and v_i correspond to the same ordinal preferences over allocations, meaning that any valid improving swap-deal for $(u_i)_{i \in \mathcal{N}}$ is also an improving swap-deal for $(v_i)_{i \in \mathcal{N}}$, and vice-versa. Since we proved the convergence to an equilibrium in $\mathcal{I}_1$ (see Proposition 8), we can also reach an equilibrium in $\mathcal{I}_2$ by applying the same sequence of swap-deals. We can conclude that the swap mechanism always converges to a stable outcome even when all agents have different utility trade-offs φ.

On the optimal social welfare. Unfortunately, not all our results can be freely translated to the case of a single φ using our transformation, as some notions such as the social welfare will be altered by the transformation. In particular, the optimal allocation of instance $\mathcal{I}_1$ may be different from the optimal allocation of the translated instance $\mathcal{I}_2$, as shown in the following example.

EXAMPLE 2. *Consider an instance $\mathcal{I}_1$ with 3 agents partitioned into 2 types $T_1 = \{1, 2\}$ and $T_2 = \{3\}$, and 3 items partitioned into 2 blocs $B_1 = \{h_1, h_2\}$ and $B_2 = \{h_3\}$. The utilities are given by:*

- $u_1^I(h_1) = 1$ *and* $u_1^I(h_2) = u_1^I(h_3) = 0$ *with* $\varphi_1 = 0.01$.
- $u_2^I(h_1) = u_2^I(h_3) = 0$ *and* $u_2^I(h_2) = 1$ *with* $\varphi_2 = 1$
- $u_3^I(h_2) = 1$ *and* $u_3^I(h_1) = u_3^I(h_3) = 0$ *with* $\varphi_3 = 0.01$.

For this instance, allocation A defined by $A(1) = h_1$, $A(2) = h_2$, and $A(3) = h_3$ is a valid allocation with maximum social welfare, with $sw(A) = 3.02$. Now let us consider the instance I_2 obtained from I_1 using our transformation. Here we have $\varphi = \min\{0.01, 1\} = 0.01$, and $(v_i)_{i \in \mathcal{N}}$ are given by:

- $v_1^I(h_1) = 1$ *and* $v_1^I(h_2) = v_1^I(h_3) = 0$.
- $v_2^I(h_1) = v_2^I(h_3) = 0$ *and* $v_2^I(h_2) = 0.01$
- $v_3^I(h_2) = 1$ *and* $v_3^I(h_1) = v_3^I(h_3) = 0$.

For instance $\mathcal{I}_2$, allocation A' defined by $A'(1) = h_1$, $A'(2) = h_3$, and $A'(3) = h_2$ is such that $sw(A') = 2.02$, whereas $sw(A) = 1.04$. Thus instances $\mathcal{I}_1$ and $\mathcal{I}_2$ have different optimal allocations.