Strategyproof Mechanisms Without Money for 2-Exchange Systems

Javier Cembrano
Department of Industrial Engineering, Universidad de Chile, Santiago, Chile

Max Klimm
Institute of Mathematics, Technische Universität Berlin, Germany

Martin Knaack
Institute of Mathematics, Technische Universität Berlin, Germany

Arturo Merino
Department of Computer Science, Universidad de Chile, Santiago, Chile

Abstract

We study a mechanism design problem in which a ground set of items E is distributed among self-interested agents. The existence of an item is private information of the agent owning it. A mechanism takes as input a set of reported items together with their intrinsic weights and returns a feasible set of items. A mechanism is strategyproof if no agent can increase the total weight of their items in the solution by withholding a subset of their items from the mechanism. It is α-approximate if the total weight of items selected is at least an $1/\alpha$-fraction of the total weight of an optimal solution.

Our main result is a 6.018-approximate strategyproof mechanism for feasibility constraints defined by a 2-exchange system. This class of independence systems is defined by a combinatorial exchange condition and includes, for example, b-matchings in general graphs, intersections of strongly base-orderable matroids, and unit interval scheduling. This constant approximation generalizes and improves over a logarithmic approximation for matchings. We also obtain a strategyproof mechanism with logarithmic approximation for generalized assignment instances, where items represent compatibilities between jobs and machines. While prior work focused on the special case in which each agent controls a single job, we provide the first logarithmic approximation guarantee for the general setting in which agents may control multiple jobs. We finally provide improved approximation guarantees for matching instances with binary weights, beating the 2-approximation given by a simple greedy mechanism for any finite number of agents and showing a strict separation between deterministic and randomized mechanisms for the case of two agents.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design

Keywords and phrases mechanism design without money, strategyproof mechanisms, approximation algorithms, 2-exchange systems, matchings

Digital Object Identifier 10.4230/LIPIcs.ESA.2026.84

Funding *Max Klimm* and *Martin Knaack*: funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, EXC-2046/2, project ID: 390685689).
Arturo Merino: supported by ANID FONDECYT Iniciación No. 11251528.

1 Introduction

Allocating items owned by strategic agents is a fundamental problem in mechanism design. One of the main challenges is that agents possess private information about the existence of items, while the mechanism itself does not. To encourage all agents to truthfully disclose all relevant information, mechanism designers seek *strategyproof* mechanisms, i.e., mechanisms in which each agent always weakly prefers the outcome resulting from reporting all items they hold over any outcome that could result from hiding a subset of these items. To facilitate the design of such mechanisms, most of the mechanism design literature allows monetary payments between the mechanism and the agents. In this context, for example, the well-known Vickrey–Clarke–Groves (VCG) mechanisms [9, 16, 25] are strategyproof and maximize social welfare in many practically relevant settings.

There are many settings, however, in which monetary payments are not feasible or undesirable. This may be due to ethical, moral, legal, or practical constraints. Notable examples include the allocation of kidney transplants [2, 22], school or residency placements [1, 14, 21], and political decision-making processes [15, 23]. These situations motivate the development of a robust theory of mechanisms without money, a research agenda put forth in a seminal paper by Procaccia and Tennenholtz [20].

In this paper, we study a general setting with a set $[n]$ of agents and a universe $U = \bigcup_{a\in[n]} U_a$ of potential items, where U_a denotes the set of items owned by agent a and each item $e \in U$ has an intrinsic weight $w_e > 0$. A *mechanism* receives a set $E = \bigcup_{a\in[n]} E_a \subseteq U$ of *reported* items with their ownership relations and their weights, and returns a *feasible* set of items. Feasibility is given by an underlying independence system $(U, \hat{\mathcal{F}})$ defined on the universe of potential items, whose restriction $(E, \hat{\mathcal{F}}|_E)$ to the set E is known to the mechanism. This is a mild assumption satisfied in many applications since it only requires that the mechanism is able to decide which subset of items reported by the agents is feasible. The utility of an agent is equal to the total weight of their items in the solution returned by the mechanism. We are interested in (randomized) mechanisms that are *strategyproof*, in the sense that no agent can increase their (expected) utility by reporting a proper subset of their items.

Among all strategyproof mechanisms, we are interested in those that provide a good approximation of the optimal weight. We call a mechanism α-approximate for some $\alpha \geq 1$ if, on any instance, the total weight of the items included in the solution computed by the mechanism is at least the total weight of the items included in an optimal solution divided by α.

1.1 Our Contribution and Techniques

We provide worst-case approximation guarantees of strategyproof mechanisms without money for the selection of items owned by strategic agents for several independence systems. Our results are summarized in Table 1.

We begin by developing a simple yet powerful technique to obtain approximation guarantees for mechanisms without money. Following an approach reminiscent of the celebrated *revelation principle* [19], we let our mechanisms internalize the best possible deviations of the agents and compensate them with higher selection probabilities for not deviating. Call a mechanism β*-strategyproof* if no agent can benefit by a factor larger than β by hiding items; then, an α-approximate and β-strategyproof mechanism can be translated into a strategyproof and $\alpha\beta$-approximate one.

This technique allows us to obtain our technically most involved result, a 6.018-approximate strategyproof mechanism for 2-exchange systems. These systems are defined via a certain exchange property and encompass, for example, b-matchings, strongly base-orderable (SBO) matroid intersection, and unit interval scheduling. Our mechanism operates by sorting the items according to class–agent pairs, where the class is given by the nearest power of a predefined base below the weight of the items after applying a random shift, and the agents have a fixed order. It first computes a lexicographically optimal feasible set according to this

Table 1 Our bounds and previous bounds on the approximation guarantees for deterministic and randomized strategyproof mechanisms in different settings.

setting	restriction	type	upper bound	lower bound
2-exchange system *(e.g., b-matching, SBO matroid intersection, unit interval scheduling)*		R	6.018 [C. 11]	$\frac{4\sqrt{2}+2}{7}$ [10]
multi-demand GAP with k jobs		R	$\mathcal{O}(\log k)$ [T. 13]	$\frac{4\sqrt{2}+2}{7}$ [10]
matching, binary weights		D	2 [P. 15]	$\frac{3}{2}$ [P. 14]
		R	$\frac{2n}{n+1}$ [T. 17]	$\frac{12}{11}$ [P. 14]
	$n = 2$	D	$\frac{3}{2}$ [T. 16]	$\frac{3}{2}$ [P. 14]

order and then subsamples the selected items. On an intuitive level, if an agent hides items from a high bucket, then the additional items they may gain in lower buckets can be charged back to the hidden ones. The subsampling step equalizes the expected contribution of items within the same bucket and removes any advantage from swapping one revealed item for a heavier revealed item from the same bucket. The positive results for b-matchings improve and generalize previous work of Chen et al. [7], who gave an $\mathcal{O}(\log m)$-approximation for matching instances with m edges.

We also obtain an $\mathcal{O}(\log k)$-approximate strategyproof mechanism for multi-demand GAP with k jobs by reducing the problem to a logarithmic number of b-matching instances, thus extending prior work beyond the setting in which each agent controls a single job [8, 10]. Our mechanism provides the first nontrivial approximation in this multi-demand setting.

Finally, for cardinality matching (i.e., binary weights), we obtain improved bounds: a deterministic tight $\frac{3}{2}$-approximation for two agents and a randomized $\frac{2n}{n+1}$-approximation for n agents. For $n = 2$, this shows a nontrivial setting where randomization provably helps to obtain better approximation guarantees. Both results are proven via charging arguments on the symmetric difference between the optimal set and the output.

1.2 Examples

Our strategyproof mechanism for 2-exchange systems can model the following applications:

b-matching: Agents propose pairwise projects between two employees, each with an associated profit. Feasibility requires that each employee v participates in at most $b(v)$ selected projects yielding a b-matching constraint. The mechanism aims to select a feasible set of projects maximizing total profit subject to these degree caps.

Intersection of two SBO matroids: A company must select a subset of jobs to run. Each job can run only on a subset of machines (known to the management), so feasible job sets correspond to the independent sets of a transversal matroid, which is strongly base-orderable [4]. Additionally, employees belong to divisions, and the company enforces per-division limits, which yields a partition matroid constraint. Hence, feasible job selections are those independent in the intersection of the transversal matroid and the partition matroid, capturing simultaneously (i) existence of a feasible job–machine matching and (ii) division quotas.

Interval scheduling: Research proposals request one unit of telescope time (as in Merrifield and Saari [18]) but specify different feasible time windows. The mechanism seeks to select a maximum-weight subset of proposals whose requested unit intervals are pairwise disjoint, i.e., an independent set in a unit interval graph.

1.3 Related Work

Mechanism design without money is a general topic discussed by Procaccia and Tennenholtz [20]. Among the models mentioned in their paper is the model by Dughmi and Ghosh [10]. In their version of the generalized assignment problem (GAP), every agent corresponds to a single vertex on the left partition of a bipartite graph. The machines on the right side have capacities, and every edge has a size and a value. Capacities, sizes, and values are common knowledge. Every agent may only report a subset of its adjacent edges to the mechanism. They gave an $\mathcal{O}(\log n)$-approximate strategyproof mechanism. They further designed strategyproof mechanisms with better approximations for special cases: a 4-approximation for the size-invariant GAP where the size of an agent does not depend on the machine it is assigned to, a 4-approximation for the value-invariant GAP where the value of an agent does not depend on the machine it is assigned to, and a 2-approximation for the multiknapsack problem where neither the size nor the value of an agent depends on the machine it is assigned to. The strategyproof $\mathcal{O}(\log n)$-approximation by Dughmi and Ghosh [10] has been improved to an $\mathcal{O}(1)$-approximation by Chen et al. [8]. Fadaei and Bichler [11] studied a different GAP model where the agents correspond to the machines whose utility is equal to the total value of items assigned to them. The private information of the agents is the set of items compatible with them. They gave a truthful 4-approximation for this model.

Chen et al. [7] studied strategyproof mechanisms for the matching problem where the graph is arbitrary and agents own arbitrary subsets of edges. For the unit-demand setting where at most one edge of each agent can be selected, they gave a 3-approximation. For the multi-demand setting, they gave an $\mathcal{O}(\log m)$-approximation where m is the number of edges. We improve this to a 6.018-approximation in the multi-demand setting. For the knapsack problem where agents own items with publicly known sizes and values, they gave a truthful $\mathcal{O}(1)$-approximate mechanism, also for the multi-demand setting. For the special case of two agents Feigenbaum and Johnson [12] gave a randomized mechanism with an approximation ratio of 1.522. For an arbitrary number of agents, Cembrano et al. [6] gave a deterministic strategyproof 3-approximation mechanism and a randomized mechanism with an approximation ratio of 2.

The implications of agents withholding information have been further studied for kidney exchange models. In these models, the vertices of the graph correspond to patient–donor pairs, and an edge between them indicates a possible cross-donation between the pairs. The agents correspond to hospitals that each control a subset of the vertices. The utility of an agent is equal to the number of its own matched pairs, where the matching is determined in two rounds. The first round matching is determined by the mechanism, and in the second, an optimal matching is determined among the hidden and other unmatched vertices. Strategyproof matching mechanisms incentivize the hospitals to never hide vertices; they have been studied extensively [2, 3, 5, 24].

2 Preliminaries

We use $\mathbb{R}_+$ to denote the non-negative reals, $\mathbb{Z}$ for the integers, and $\mathbb{N}$ for the natural numbers. For $n \in \mathbb{N}$, we let $[n] \coloneqq \{1, \ldots, n\}$ denote the set of the first n natural numbers. An *independence system* is a pair $(E, \mathcal{F})$, where E is a ground set of items with $|E| = m$ and $\mathcal{F} \subseteq 2^E$ is a set system satisfying the following properties: *(i)* $\emptyset \in \mathcal{F}$; and *(ii)* $S \subset T \subseteq E$ and $T \in \mathcal{F}$ imply $S \in \mathcal{F}$. For $E' \subseteq E$, the restriction of $\mathcal{F}$ to E' is $\mathcal{F}|_{E'} \coloneqq \{S \cap E' \mid S \in \mathcal{F}\}$. We will refer to the subsets of E that belong to $\mathcal{F}$ as *independent sets* or *feasible sets*; the remaining subsets of E are called *dependent sets* or *infeasible sets*. A set $S \subseteq E$ is called a *base* if it is an inclusion-wise maximal independent set.

Next, we formalize our notion of strategyproof mechanisms for general independence systems and define 2-exchange systems, a class of independence systems introduced by Feldman et al. [13].

2.1 Strategyproof Mechanisms

We consider a universe U of potential items to report with an associated independence system $(U, \hat{\mathcal{F}})$. Each item $e \in U$ has an intrinsic weight $w_e > 0$. The set of items $U_a \subseteq U$ controlled by each agent $a \in [n]$ is private information of agent a (unknown to the mechanism and to other agents). We consider items as being unique to their owner; effectively, if two agents possess identical items, they are treated as distinct elements in U. This ensures that the sets $\{U_a\}_{a \in [n]}$ are disjoint by definition. Each agent reports a (potentially strict) subset of items $E_a \subseteq U_a$, and while doing so, they reveal the intrinsic weights of their items $w_e > 0$ for $e \in E_a$. We use $E = \bigcup_{a \in [n]} E_a$ to denote the items reported by the agents.

A *mechanism* is a (possibly random) function that receives the reported items in E with their ownership relations, the restricted independence family $\mathcal{F} = \hat{\mathcal{F}}|_E$, and the weights of items in E, and outputs an independent set $S \in \mathcal{F}$. We allow w to be a weight function on a superset of E and interpret it as the restriction of w on E. We compress the relevant information available to the mechanism in a tuple $(E, \mathcal{F}, w)$, where the partition $E = \bigcup_{a \in [n]} E_a$ is implicit. We often refer to such a tuple as an *instance* and denote it by $\mathcal{I}$. Since $\mathcal{F}$ is uniquely determined by E, we often denote the input of a mechanism simply by the pair (E, w).

We write $S_a \coloneqq S \cap E_a$ for any $S \subseteq E$ and $a \in [n]$. When needed, we denote the feasible sets of an instance $\mathcal{I}$ by $\mathcal{F}(\mathcal{I})$. For $S \subseteq \bigcup_{a \in [n]} E_a$, we let $w(S) \coloneqq \sum_{e \in S} w_e$ denote the total weight of S and $w_a(S) \coloneqq \sum_{e \in S_a} w_e$ denote the weight of the items of agent a in set S.

For an instance $(E, \mathcal{F}, w)$ and an agent $a \in [n]$, a *feasible deviation* of agent a is a reported set $E' \subseteq E$ obtained by hiding only items of agent a; equivalently, $E \setminus E' \subseteq E_a$. A deterministic mechanism M is *strategyproof* on a certain set of instances if, for every instance $(E, \mathcal{F}, w)$ in this set, every $a \in [n]$, and every feasible deviation E' of agent a, we have $w_a(\mathrm{M}(E, w)) \geq w_a(\mathrm{M}(E', w))$. A randomized mechanism M is *strategyproof* on a set of instances if, for every instance $(E, \mathcal{F}, w)$ in this set, every $a \in [n]$, and every feasible deviation E' of agent a, we have $\mathbb{E}[w_a(\mathrm{M}(E, w))] \geq \mathbb{E}[w_a(\mathrm{M}(E', w))]$. We also let $\textsc{Opt}(E, w) \coloneqq \arg\max\{w(S) \mid S \in \mathcal{F}\}$ denote the optimal feasible set for an instance.[1] A mechanism M is α-approximate on a certain set of instances if, for every instance $(E, \mathcal{F}, w)$ in this set, it provides an α-approximation of the optimal weight (possibly in expectation), i.e., $\alpha \cdot \mathbb{E}[w(\mathrm{M}(E, w))] \geq w(\textsc{Opt}(E, w))$.

[1] Here and throughout the paper, we assume a consistent tie-breaking rule. As a consequence $\max\{w(S) \mid S \in \mathcal{F}\}$ is achieved by a unique set. This can be implemented, e.g., by perturbing weights according to the edge labels.

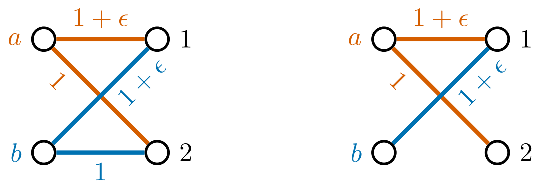

Figure 1 Example implying a lower bound of 2 on the approximation of the social welfare of deterministic strategyproof mechanisms. In each graph, the vertices on the left are the agents and the vertices on the right are the machines. All sizes and capacities are 1; the weights of the edges are written next to them.

The following lemma extends that of Cembrano et al. [6, Lemma 2.1] for the withholding setting under knapsack constraints, and is useful to prove the strategyproofness of our mechanisms. It states that it suffices to consider agents hiding a single item to prove the strategyproofness of a mechanism on a down-closed class of instances, i.e., a class of instances such that removing elements from the ground set keeps the instance within the class. While we state it for randomized mechanisms, it naturally has the same implication for deterministic mechanisms. The proof is deferred to the full version.

▶ **Lemma 1.** *Let* M *be a mechanism that is not strategyproof on a down-closed class of instances. Then, there exists an instance* $(E, \mathcal{F}, w)$ *in this set, an agent* $a \in [n]$*, and an item* $f \in E_a$ *such that* $\mathbb{E}[w_a(\mathrm{M}(E, w))] < \mathbb{E}[w_a(\mathrm{M}(E \setminus \{f\}, w))]$.

To illustrate the intrinsic challenges of designing strategyproof mechanisms, we recall in Figure 1 the example given by Dughmi and Ghosh [10, Theorem 3.3] in the context of GAP, where a mechanism has to assign agents (vertices on the left) to machines (vertices on the right). In the graph on the left, a mechanism needs to assign agents a and b to a machine each to be 2-approximate. As the situation is symmetric, it is without loss of generality to assume that agent a is assigned to 1 and agent b is assigned to 2. In the graph on the right, agent b hides edge $\{b, 2\}$. Due to strategyproofness, in this graph agent b cannot be assigned to 1, so no deterministic strategyproof mechanism can be better than 2-approximate.

2.2 2-Exchange Systems

An independence system $(E, \mathcal{F})$ is a *2-exchange system* if for every pair of independent sets $I, I' \in \mathcal{F}$ there is a bipartite graph $\mathcal{G}(I, I')$ called the *exchange graph* with the following properties:

(i) The two partition classes of vertices are $I \setminus I'$ and $I' \setminus I$;
(ii) for every $S \subseteq I' \setminus I$, we have that $I \cup S \setminus N(S) \in \mathcal{F}$;
(iii) the maximum degree of $\mathcal{G}(I, I')$ is at most 2;

where $N(S)$ denotes the set of neighbors of vertices in S.

While the definition of 2-exchange systems may look technical, it is satisfied by natural systems such as b-matchings, intersections of two strongly base-orderable matroids, and unit interval scheduling; see the full version for more details. We refer to instances of the form $(E, \mathcal{F}, w)$, where $(E, \mathcal{F})$ is a 2-exchange system, as *2-exchange instances*.

We will make extensive use of the fact that the definition of 2-exchange systems can be made two-sided.

▶ **Lemma 2.** *Let* $(E, \mathcal{F})$ *be a 2-exchange system. Then, for every* $I, I' \in \mathcal{F}$*, the exchange graph* $\mathcal{G}(I, I')$ *is such that, for every* $S \subseteq I \setminus I'$*, we have that* $I' \cup S \setminus N(S) \in \mathcal{F}$*.*

The proof is deferred to the full version. We also use the following approximation guarantee of Greedy on 2-exchange systems, as noted by Feldman et al. [13]. Here, Greedy starts with the empty set and the items in non-increasing order by weight, and iteratively adds the item to the current set if it maintains feasibility.

▶ **Lemma 3.** *Let $(E, \mathcal{F}, w)$ be a 2-exchange instance and let G be the output of* Greedy*. Then $w(G) \geq \frac{1}{2} w(\text{Opt}(E, w))$, independently of the non-increasing order used by* Greedy*.*

3 From approximate strategyproofness to strategyproofness

In this section, we introduce and apply a reduction technique from approximate strategyproofness to strategyproofness. In a nutshell, the idea is to internalize in the mechanism the agents' best possible deviations. We show that, if these deviations allow agents to improve their weight by a factor of at most β in the original mechanism, scaling down the selection probabilities by at most β suffices to incentivize truthful reporting. Thus, we effectively translate an approximately strategyproof mechanism into a strategyproof one, at the cost of an additional loss in the approximation factor.

Formally, call a mechanism *β-strategyproof* on a certain set of instances, for $\beta \geq 1$, if for every instance $(E, \mathcal{F}, w)$ in this set, every agent $a \in [n]$, and every feasible deviation E' of agent a, the inequality $\beta \cdot \mathbb{E}[w_a(\mathrm{M}(E, w))] \geq \mathbb{E}[w_a(\mathrm{M}(E', w))]$ holds. The following lemma formalizes our reduction.

▶ **Lemma 4.** *If there is a mechanism* M *that is α-approximate and β-strategyproof on a set of instances, then there is a mechanism* M′ *that is strategyproof and $\alpha\beta$-approximate on this set of instances.*

Proof. Let M be as in the statement and let $(E, \mathcal{F}, w)$ be an arbitrary instance of the considered set. We first describe M′ as a probability distribution over the elements. For every $a \in [n]$ and $e \in E_a$, if $\mathbb{P}[e \in \mathrm{M}(E, w)] = 0$ we fix $\mathbb{P}[e \in \mathrm{M}'(E, w)] = 0$, and if $\mathbb{P}[e \in \mathrm{M}(E, w)] > 0$ we fix

$$\mathbb{P}[e \in \mathrm{M}'(E, w)] = \frac{\max\{\mathbb{E}[w_a(\mathrm{M}(E^*, w))] \mid E^* \subseteq E \text{ and } E \setminus E^* \subseteq E_a\}}{\beta \cdot \mathbb{E}[w_a(\mathrm{M}(E, w))]} \cdot \mathbb{P}[e \in \mathrm{M}(E, w)].$$

The latter is well defined, since $\mathbb{P}[e \in \mathrm{M}(E, w)] > 0$ with $w_e > 0$ implies $\mathbb{E}[w_a(\mathrm{M}(E, w))] > 0$. Note that we have for every $e \in E$ that $\mathbb{P}[e \in \mathrm{M}'(E, w)] \leq \mathbb{P}[e \in \mathrm{M}(E, w)]$, since M is β-strategyproof. We use this fact to prove that M′ corresponds to a distribution over independent sets; the proof is deferred to the full version.

▷ Claim 5. M′ can be realized as a distribution over independent sets.

The approximation factor of M′ holds because M is assumed to be α-approximate, and since we have for every $e \in E$ that $\beta \cdot \mathbb{P}[e \in \mathrm{M}'(E, w)] \geq \mathbb{P}[e \in \mathrm{M}(E, w)]$. Finally, to see that M′ is strategyproof, consider an agent $a \in [n]$ and a feasible deviation E' of agent a. The case where $\mathbb{E}[w_a(\mathrm{M}(E', w))] = 0$ is clear, and if $\mathbb{E}[w_a(\mathrm{M}(E', w))] > 0$, we have that

$$\begin{aligned}
&\mathbb{E}[w_a(\mathrm{M}'(E', w))] \\
&= \sum_{e \in E'_a} \frac{\max\{\mathbb{E}[w_a(\mathrm{M}(E^*, w))] \mid E^* \subseteq E' \text{ and } E' \setminus E^* \subseteq E'_a\}}{\beta \cdot \mathbb{E}[w_a(\mathrm{M}(E', w))]} \cdot \mathbb{P}[e \in \mathrm{M}(E', w)] \cdot w_e \\
&= \frac{1}{\beta} \max\{\mathbb{E}[w_a(\mathrm{M}(E^*, w))] \mid E^* \subseteq E' \text{ and } E' \setminus E^* \subseteq E'_a\},
\end{aligned}$$

which is maximized by reporting all items E_a. ◀

3.1 The bucket-quota mechanism and lexicographic optimization

We introduce a natural mechanism that we will show to be approximately strategyproof and approximately optimal on 2-exchange systems, allowing us to apply our reduction from Lemma 4. The mechanism groups items into class–agent buckets according to rounded weights after a random shift, computes a lex-optimal feasible set with respect to these buckets, and then subsamples within each bucket.

We introduce some notation to define the mechanism formally. Given $b > 1$ and an instance $(E, \mathcal{F}, w)$, we define the *class–agent buckets* for a shift $x \in [0, 1]$ as follows:

$$B_{i,a}(x) = \left\{ e \in E_a \;\middle|\; w_e \in [b^{x+i}, b^{x+i+1}) \right\} \quad \text{for every } a \in [n],\ i \in \mathbb{Z}.$$

Whenever x is fixed, we omit it from this notation.

We now define various orders over class–agent pairs and items. For class–agent pairs (i, a) and (j, b), we write $(i, a) \prec (j, b)$ if either $i < j$ or $i = j$ and $a < b$. As usual, we use the notation $(i, a) \preceq (j, b)$ to mean that either $(i, a) \prec (j, b)$ or $(i, a) = (j, b)$. Naturally, this induces an order over items given by the ordering of their buckets; i.e., for $x \in [0, 1]$ and $e, f \in E$, we say that $e \prec f$ if there are two class–agent pairs (i, a) and (j, b) such that $e \in B_{i,a}(x)$ and $f \in B_{j,b}(x)$ with $(i, a) \prec (j, b)$.

We now extend this order to sets of items. For a fixed $x \in [0, 1]$, we say that a set $S \subseteq E$ is *(class–agent) lex-larger* than $S' \subseteq E$ for the given instance if there is a class–agent pair (i, a) such that $|S \cap B_{i,a}(x)| > |S' \cap B_{i,a}(x)|$ and for all (j, b) with $(j, b) \succ (i, a)$ we have $|S \cap B_{j,b}(x)| = |S' \cap B_{j,b}(x)|$. Similarly, we say that two sets $S, S' \subseteq E$ are *(class–agent) lex-equal* if for all class–agent pairs (i, a), we have $|S \cap B_{i,a}(x)| = |S' \cap B_{i,a}(x)|$. We overload the $\prec$ notation and use it also for sets; i.e., for sets $S, S' \subseteq E$, we use the notations $S' \prec S$, $S' \preceq S$ and $S' \sim S$ to mean that S is lex-larger than S', lex-larger or lex-equal to S', and lex-equal to S', respectively.

Now we are ready to define the BucketQuota mechanism, which is formally described in Algorithm 1. The mechanism first chooses a random shift x and computes a lex-optimal independent set M^* according to the class–agent buckets defined by x. Then, for each item in M^*, it assigns a selection probability that is proportional to the lower endpoint of the weight interval of its bucket and inversely proportional to its weight, and adds it to the output with this probability. Note that the lex-optimal independent set can be reduced to a standard maximum-weight optimization by assigning exponentially separated surrogate weights to the buckets, so it is tractable whenever the underlying optimization problem is tractable.

3.2 A strategyproof mechanism for 2-exchange systems

In this section, we analyze BucketQuota and show that it is approximately strategyproof and approximately optimal on 2-exchange instances. Using Lemma 4, this yields a mechanism that is strategyproof and approximately optimal.

The next lemma states the approximate strategyproofness of BucketQuota on 2-exchange instances.

▶ **Lemma 6.** *For $b > 1$, BucketQuota with base b is $\frac{b+1}{b-1}$-strategyproof on 2-exchange instances.*

To prove this lemma, we compare the lex-optimal candidate sets obtained from the truthful report and from a report where a single agent hides items. First, we note that the subsampling makes every item in the same bucket contribute the same expected weight,

Algorithm 1 BucketQuota with base $b > 1$.

Input: An instance $(E, \mathcal{F}, w)$.
Output: A set $M \in \mathcal{F}$.
Choose $x \in [0,1]$ uniformly at random.
$M^* \leftarrow$ a lex-optimal independent set for the class–agent buckets $\{B_{i,a}(x)\}_{i\in\mathbb{Z},a\in[n]}$.
$M \leftarrow \emptyset$.
for $e \in M^*$ **do**
  $p_e \leftarrow b^{x+i}/w_e$, where (i,a) is such that $e \in B_{i,a}(x)$.
  Add e to M with probability p_e.
return M

essentially reducing to the cardinality case. We then consider a charging argument. Here, we restrict to one connected component of the exchange graph, and show that any extra item gained by the agent in a bucket $B_{i,a}$ can be charged to a hidden item of that agent in a higher bucket. Then, we use properties of the exchange graph to show that every hidden item gets charged not too often, which ultimately allows us to bound the improvement of the agent.

Proof of Lemma 6. We consider an instance $(E, \mathcal{F}, w)$, where $(E, \mathcal{F})$ is a 2-exchange system. We fix a realization of $x \in [0,1]$ for the whole proof; we thus omit this argument from the class–agent buckets notation. Let $a \in [n]$ and let $E' \subseteq E$ be such that $E \setminus E' \subseteq E_a$, that is, only items of agent a may be hidden. We abbreviate the outcome of BucketQuota for (E, w) and (E', w) by $\textsc{Alg}(E, w)$ and $\textsc{Alg}(E', w)$, respectively.

We consider the class–agent buckets $\{B_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$ and $\{B'_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$ corresponding to the reported sets E and E', respectively. Let $I \in \mathcal{F}$ be a lex-optimal independent set for the class–agent buckets $\{B_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$, and let $I' \in \mathcal{F}|_{E'}$ be a lex-optimal independent set for the class–agent buckets $\{B'_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$. We differentiate between the orders induced by the class–agent buckets $\{B_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$ and $\{B'_{i,b}\}_{i\in\mathbb{Z},b\in[n]}$, denoting them by $\prec$ and $\prec'$, respectively. Let S be the set of items in I that agent a hides, i.e., $S \coloneqq I \cap (E_a \setminus E')$.

We now consider $\mathcal{G} = \mathcal{G}(I, I')$, the bipartite graph given by the definition of a 2-exchange system for $I, I' \in \mathcal{F}$. As in that definition, for a set $T \subseteq I \Delta I'$ we denote its set of neighbors in this graph by $N(T)$. For $T \subseteq I\Delta I'$ and $i \in \mathbb{Z}$, we define $d_i(T)$ as the number of elements gained by agent a in the bucket $B_{i,a}$ restricted to T; i.e., $d_i(T) \coloneqq |T \cap I' \cap B_{i,a}| - |T \cap I \cap B_{i,a}|$. We claim the following.

▷ Claim 7. For every connected component of $\mathcal{G}$, with vertex set F, and every $i \in \mathbb{Z}$, we have that $\sum_{j>i} |F \cap S \cap B_{j,a}| \geq \frac{d_i(F)}{2}$.

Before proving the claim, we show how it allows us to conclude the lemma. Let p_e be as in Algorithm 1. Then, we have

$$\begin{aligned}
\mathbb{E}[w_a(\textsc{Alg}(E', w))] - \mathbb{E}[w_a(\textsc{Alg}(E, w))] &= \sum_{e\in I'\cap E'_a} p_e w_e - \sum_{e\in I\cap E_a} p_e w_e \\
&= \sum_{i\in\mathbb{Z}} (|I' \cap B_{i,a}| - |I \cap B_{i,a}|) b^{x+i} \\
&= \sum_{i\in\mathbb{Z}} \sum_{F} d_i(F) b^{x+i},
\end{aligned}$$

where the sum over F is over the vertex sets of connected components of $\mathcal{G}$. Using Claim 7 we obtain

$$\begin{aligned}\sum_{i\in\mathbb{Z}}\sum_{F} d_i(F)b^{x+i} &\leq 2\sum_{i\in\mathbb{Z}}\sum_{F}\sum_{j>i}|F\cap S\cap B_{j,a}|b^{x+i}\\ &= 2\sum_{i\in\mathbb{Z}}\sum_{j>i}|B_{j,a}\cap S|b^{x+i},\end{aligned}$$

where the last equality uses that the sets F partition the vertex set. Exchanging sums, we obtain that

$$\begin{aligned}2\sum_{i\in\mathbb{Z}}\sum_{j>i}|B_{j,a}\cap S|b^{x+i} &= 2\sum_{j\in\mathbb{Z}}|B_{j,a}\cap S|\sum_{i<j}b^{x+i}\\ &= \frac{2}{b-1}\sum_{j\in\mathbb{Z}}|B_{j,a}\cap S|b^{x+j}\\ &= \frac{2}{b-1}\sum_{e\in S}p_e w_e\\ &\leq \frac{2}{b-1}\mathbb{E}[w_a(\textsc{Alg}(E,w))],\end{aligned}$$

from which we conclude that $\mathbb{E}[w_a(\textsc{Alg}(E',w))] \leq \frac{b+1}{b-1}\mathbb{E}[w_a(\textsc{Alg}(E,w))]$.

We now prove Claim 7.

Given F and $i \in \mathbb{Z}$ as in the statement, we say that a set of paths $\mathcal{P} = \{P_1, \dots, P_{d_i(F)}\}$ in F is an i-covering of F if

- **(i)** each item in F belongs to at most two paths in $\mathcal{P}$;
- **(ii)** the neighbors of the extreme points of each path P_ℓ that are not in P_ℓ lie in $I' \cap B_{i,a}$, i.e., $N(P_\ell) \setminus P_\ell \subseteq I' \cap B_{i,a}$ for $\ell \in [d_i(F)]$;
- **(iii)** $d_i(P_\ell) = 1$ for each $\ell \in [d_i(F)]$.

▷ Claim 8. If $d_i(F) > 0$, there exists an i-covering of F.

The proof of Claim 8 is deferred to the full version; we now use it to conclude Claim 7.

If $d_i(F) = 0$ the claim follows trivially, so we assume that $d_i(F) > 0$ in what follows. Let $\mathcal{P} = \{P_1, \dots, P_{d_i(F)}\}$ be an i-covering of F. We fix $\ell \in [d_i(F)]$ and denote $P = P_\ell$.

We aim to charge the items gained in the path P to an item in P that agent a hides and loses from a higher bucket, i.e., an item in $P \cap S \cap B_{j,a}$ for some $j > i$. To this end, we suppose for the sake of contradiction that there is no such item; i.e.,

$$P \cap S \cap B_{j,a} = \emptyset \text{ for every } j > i. \tag{1}$$

We begin with the following observation: There exists $(i^*, a^*) \succ (i, a)$ such that

$$|P\cap I\cap B_{i^*,a^*}| > |P\cap I'\cap B_{i^*,a^*}| \text{ and } |P\cap I\cap B_{j,b}| = |P\cap I'\cap B_{j,b}| \text{ for every } (j,b) \succ (i^*,a^*). \tag{2}$$

Otherwise, property *(ii)* in the definition of i-covering would yield $I \prec I\Delta P$, and

$$I\Delta P = I \cup (P\cap I') \setminus (P \cap I) = I \cup (P \cap I') \setminus N(P\cap I') \in \mathcal{F},$$

since $P \cap I = N(P \cap I')$ from property *(ii)* in the definition of i-covering. Hence, this contradicts the lex-optimality of I.

We now claim that

$$P \cap B_{j,b} = P \cap B'_{j,b} \text{ for every } (j,b) \succ (i,a). \tag{3}$$

Indeed, if $b \neq a$, then agent a is the only one who can hide items, so $B_{j,b} = B'_{j,b}$. If $b = a$ and $(j,a) \succ (i,a)$, then necessarily $j > i$, and (1) implies that P contains no hidden item of agent a in bucket $B_{j,a}$.

We use (2) and (3) to conclude that

$$|P \cap I \cap B'_{i^*,a^*}| \overset{(3)}{=} |P \cap I \cap B_{i^*,a^*}| \overset{(2)}{>} |P \cap I' \cap B_{i^*,a^*}| \overset{(3)}{=} |P \cap I' \cap B'_{i^*,a^*}|,$$

and for every $(j,b) \succ (i^*,a^*)$ that

$$|P \cap I \cap B'_{j,b}| \overset{(3)}{=} |P \cap I \cap B_{j,b}| \overset{(2)}{=} |P \cap I' \cap B_{j,b}| \overset{(3)}{=} |P \cap I' \cap B'_{j,b}|.$$

But then, $I' \prec' I' \cup (P \cap I) \setminus N(P \cap I)$ as the neighbors of the endpoints of P that are not in P are contained in $B_{i,a}$ from property *(ii)* in the definition of i-covering. Since $I' \cup (P \cap I) \setminus N(P \cap I) \in \mathcal{F}$, this contradicts the lex-optimality of I' with respect to the order $\prec'$.

We conclude that $P \cap S \cap B_{j,a} \neq \emptyset$ for some $j > i$. Thus, for each $\ell \in [d_i(F)]$ we can charge the path P_ℓ to the corresponding item. Since each item belongs to at most two paths due to property *(i)* in the definition of i-covering, we conclude that

$$\sum_{j>i} |F \cap S \cap B_{j,a}| \geq \frac{1}{2} \sum_{j>i} \sum_{\ell \in [d_i(F)]} |P_\ell \cap S \cap B_{j,a}| \geq \frac{d_i(F)}{2}. \qquad \blacktriangleleft$$

The same mechanism also preserves, up to the expected loss from bucketing ($\int_0^1 b^{-x} dx = \frac{b-1}{b \ln b}$), the constant-factor approximation guarantee of Greedy on 2-exchange systems.

▶ **Lemma 9.** *For $b > 1$, BucketQuota with base b is $\frac{2b \ln b}{b-1}$-approximate on 2-exchange instances.*

This analysis is inspired by the work of Matuschke et al. [17] and is deferred to the full version.

Combining Lemmas 4, 6, and 9, we directly obtain the main result of this section.

▶ **Theorem 10.** *For every $b > 1$, there is a mechanism that is strategyproof and $\frac{2b(b+1)\ln b}{(b-1)^2}$-approximate on 2-exchange instances.*

In particular, choosing $b = 5.7$ yields a 6.018-approximate mechanism.

▶ **Corollary 11.** *There is a mechanism that is strategyproof and 6.018-approximate on 2-exchange instances.*

3.3 Applications of 2-exchange systems

Standard subclasses of 2-exchange systems include b-matchings, intersections of two strongly base-orderable matroids, SBO matchoids, SBO matroid parity systems, and unit interval scheduling; see the full version for the corresponding definitions and background. Thus, Corollary 11 immediately yields the following.

▶ **Corollary 12.** *There is a mechanism that is strategyproof and 6.018-approximate on b-matching instances, SBO matroid intersection instances, SBO matchoid instances, SBO matroid parity instances, and unit interval scheduling instances.*

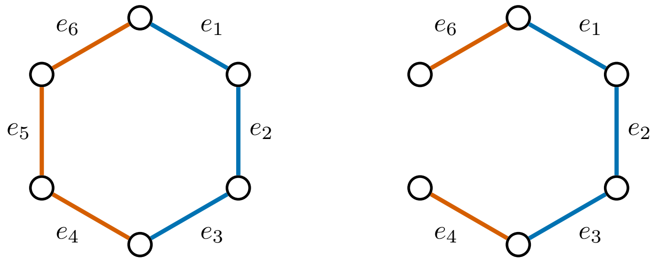

Figure 2 Instances used to prove the lower bounds on the approximation factor attainable by strategyproof mechanisms on cardinality matching instances. Edges of the same color belong to the same agent. For the deterministic bound, we observe that a mechanism must select three edges on the instance at the left to beat $\frac{3}{2}$-approximation and assume without loss of generality that it selects $\{e_1, e_3, e_5\}$. In the instance at the right, it must select $\{e_2, e_4, e_6\}$, so the orange agent benefited from hiding e_5. The proof is similar for randomized mechanisms.

3.4 Multi-demand generalized assignment problem

We additionally consider multi-demand generalized assignment instances, where agents own arbitrary subsets of job–machine edges in a bipartite graph with public job sizes and machine capacities. By rounding sizes and reducing to a logarithmic number of b-matching instances, we obtain the following result; the formal setting and construction are deferred to the full version.

▶ **Theorem 13.** *There is a mechanism that is strategyproof and $\mathcal{O}(\log k)$-approximate on multi-demand generalized assignment instances with k jobs.*

4 Cardinality matching

In this section, we study a restricted version of the matching setting from the previous section, where all edges have intrinsic unit weight. We give strategyproof mechanisms achieving better approximation guarantees and show a strict separation for two agents between the best-possible approximation that deterministic and randomized mechanisms can achieve.

In *cardinality matching* instances, the ground set E corresponds to the set of edges of a simple, undirected graph $G = (V, E)$, and the independent sets correspond to subsets of edges with no common endpoints; i.e., $S \subseteq E$ is an independent set if $e \cap f = \emptyset$ for every $e, f \in S$ with $e \neq f$. The weight of a set of edges corresponds to its cardinality, i.e., $w_e = 1$ for every $e \in E$. Therefore, throughout this section, we only use the set of reported edges E as the input of a mechanism, writing $\mathrm{M}(E)$ and $\textsc{Opt}(E)$ for the outcome of a mechanism M and for the optimal set, respectively. An instance within this class is fully represented by the graph $G = (V, E)$; we thus refer to the set of matchings for a given graph by $\mathcal{F}(G)$. Lemma 1 extends immediately to this case, stating that a mechanism M is strategyproof if we have $|\mathrm{M}_a(E)| \geq |\mathrm{M}_a(E \setminus \{e\})|$ for every set of items E, agent a, and $e \in E_a$.

We first state lower bounds on the approximation guarantees that deterministic and randomized strategyproof mechanisms can achieve in this setting. The proof is deferred to the full version; the instances we use in it are shown in Figure 2.

▶ **Proposition 14.** *Let M be a mechanism that is strategyproof and α-approximate on cardinality matching instances. Then, $\alpha \geq \frac{12}{11} \approx 1.0909$. If M is deterministic, $\alpha \geq \frac{3}{2}$.*

On the upper-bound side, a simple greedy baseline we call GreedyByAgent already yields the following guarantee; the mechanism and proof are deferred to the full version. The result is stated in the more general context of k-system instances, i.e., instances $(E, \mathcal{F}, w)$ where, for each subset $S \subseteq E$, the bases of S in the system $(E, \mathcal{F})$ do not differ in size by a factor greater than k.

▶ **Proposition 15.** GreedyByAgent is strategyproof and k-approximate on k-system instances with binary weights.

We next improve the deterministic guarantee for two agents.

▶ **Theorem 16.** *There exists a deterministic mechanism that is strategyproof and $\frac{3}{2}$-approximate on cardinality matching instances with $n = 2$ agents.*

The mechanism reserves for one agent the size of a maximum individual matching and uses the choice among such matchings to benefit the other agent as much as possible. This quota-and-tie-breaking idea avoids the bad 3-path configurations behind the 2-approximation barrier; the formal mechanism and proof are deferred to the full version.

A randomized variant of this mechanism chooses the privileged agent uniformly at random and achieves the ratio $\frac{2n}{n+1}$.

▶ **Theorem 17.** *There exists a strategyproof and $\frac{2n}{n+1}$-approximate mechanism on cardinality matching instances with n agents.*

The description of the mechanism, the proof of the theorem, and a discussion about the tightness of this analysis are deferred to the full version.

5 Concluding Remarks

In this paper, we studied mechanisms without money for selecting subsets of items owned by strategic agents in combinatorial domains captured by 2-exchange systems. We provided strategyproof mechanisms with improved approximation guarantees for settings that generalize previously studied domains by allowing arbitrary item ownership.

Our main results are driven by a simple reduction that transforms approximately strategyproof mechanisms, where agents can gain by at most a bounded factor from hiding items, into fully strategyproof mechanisms while losing only the same factor in approximation. Combined with bucketing, lexicographic optimization, and subsampling, this yields a constant-factor mechanism for a broad class of systems, including b-matchings, intersections of SBO matroids, and unit interval scheduling. We also obtained a logarithmic approximation for multi-demand GAP with arbitrary item ownership and refined the picture for cardinality matching, improving on the simple greedy 2-approximation and exhibiting a strict separation between deterministic and randomized mechanisms for two agents.

Beyond tightening the gaps that remain, we believe that the reduction to strategyproofness, together with bucketing and subsampling, is interesting on its own and deserves further exploration in other mechanism design settings. Natural next steps include broader assignment-type problems, 2-systems, and general independence systems with binary weights. Another natural direction is to strengthen the model to allow further deviations by the agents; e.g., by allowing them to underreport weights.

References

1 Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93:729–747, 2003. doi:10.1257/000282803322157061.

2 Itai Ashlagi, Felix A. Fischer, Ian A. Kash, and Ariel D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 91:284–296, 2015. doi:10.1016/J.GEB.2013.05.008.

3 Itai Ashlagi and Alvin E. Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9:817–863, 2014. doi:10.3982/TE1357.

4 Richard A. Brualdi and Edward B. Scrimger. Exchange systems, matchings, and transversals. *J. Combin. Theory*, 5:244–257, 1968. doi:10.1016/S0021-9800(68)80071-7.

5 Ioannis Caragiannis, Aris Filos-Ratsikas, and Ariel D. Procaccia. An improved 2-agent kidney exchange mechanism. *Theoretical Computer Science*, 589:53–60, 2015. doi:10.1016/J.TCS.2015.04.013.

6 Javier Cembrano, Max Klimm, and Martin Knaack. Packing a knapsack with items owned by strategic agents. In *Proceedings of the 20th Conference on Web and Internet Economics*, 2024.

7 Ning Chen, Nick Gravin, and Pinyan Lu. Mechanism design without money via stable matching, 2011. arXiv:1104.2872.

8 Ning Chen, Nick Gravin, and Pinyan Lu. Truthful generalized assignments via stable matching. *Mathematics of Operations Research*, 39(3):722–736, 2014. doi:10.1287/MOOR.2013.0625.

9 Edward H. Clarke. Multipart pricing of public goods. *Public Choice*, 11:17–33, 1971. doi:10.1007/bf01726210.

10 Shaddin Dughmi and Arpita Ghosh. Truthful assignment without money. In David C. Parkes, Chrysanthos Dellarocas, and Moshe Tennenholtz, editors, *Proceedings of the 11th ACM Conference on Electronic Commerce*, pages 325–334, 2010. doi:10.1145/1807342.1807394.

11 Salman Fadaei and Martin Bichler. Generalized assignment problem: Truthful mechanism design without money. *Operations Research Letters*, 45(1):72–76, 2017. doi:10.1016/J.ORL.2016.12.003.

12 Itai Feigenbaum and Matthew P. Johnson. Selfish knapsack. In Satinder Singh and Shaul Markovitch, editors, *Proceedings of the 31st Conference on Artificial Intelligence*, pages 509–515, 2017. doi:10.1609/AAAI.V31I1.10579.

13 Moran Feldman, Joseph Naor, Roy Schwartz, and Justin Ward. Improved approximations for k-exchange systems. In *Proceedings of the 19th Annual European Symposium of Algorithms*, pages 784–798. Springer, Heidelberg, 2011. doi:10.1007/978-3-642-23719-5_66.

14 David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69:9–15, 1962. doi:10.2307/2312726.

15 Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973. doi:10.2307/1914083.

16 Theodore Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973. doi:10.2307/1914085.

17 Jannik Matuschke, Martin Skutella, and José A Soto. Robust randomized matchings. *Mathematics of Operations Research*, 43(2):675–692, 2018. doi:10.1287/moor.2017.0878.

18 Michael R. Merrifield and Donald G. Saari. Telescope time without tears: A distributed approach to peer review. *Astronomy & Geophysics*, 50:16–20, 2009.

19 Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. doi:10.1287/moor.6.1.58.

20 Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4):18:1–18:26, 2013. doi:10.1145/2542174.2542175.

21 Alvin E. Roth. The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92, 1984. doi:10.1086/261272.

22 Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *American Econonomic Review*, 97:828–851, 2007. doi:10.1257/aer.97.3.828.

23 Mark A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econ. Theory*, 10:187–217, 1975.

24 Panagiotis Toulis and David C. Parkes. A random graph model of kidney exchanges: Efficiency, individual-rationality and incentives. In Yoav Shoham, Yan Chen, and Tim Roughgarden, editors, *Proceedings of the 12th ACM Conference on Electronic Commerce*, pages 323–332, 2011. doi:10.1145/1993574.1993626.

25 William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16:8–37, 1961. doi:10.1111/j.1540-6261.1961.tb02789.x.