Rank-optimal assignments in uniform markets

Afshin Nikzad*

March 17, 2021

Abstract

We prove that in a market where agents rank objects independently and uniformly at random, there exists an assignment of objects to agents with a constant average rank (i.e., an average rank independent of the market size). The proof builds on techniques from random graph theory and the FKG inequality (Fortuin et al., 1971). When the agents' rankings are their private information, no Dominant Strategy Incentive Compatible mechanism can implement the assignment with the smallest average rank; however, we show that there exists a Bayesian Incentive Compatible mechanism that does so. Together with the fact that the average rank under the Random Serial Dictatorship (RSD) mechanism grows infinitely large with the market size, our findings indicate that the average rank under RSD can take a heavy toll compared to the first-best, and highlight the possibility of using other assignment methods in scenarios where average rank is a relevant objective.

Keywords: Matching, average rank, random serial dictatorship, FKG inequality

JEL Codes: C78, D82

*Department of Economics, USC. Email: afshin.nikzad@usc.edu

1 Introduction

In assignment problems involving ordinal preferences, *rank* is a commonly considered notion for comparing different assignments. School districts consider the number of students assigned to their first choice, second choice, and so on, to report their assignment results or compare different assignments (Abdulkadirouglu et al., 2005; De Haan et al., 2015; Featherstone, 2020). Similarly, some institutions such as Teach For America choose assignments explicitly based on such rank distributions (Featherstone, 2020).[1]

The notion of rank is also used in evaluating assignments in the National Residency Matching Program (NRMP, 2020).[2] In evaluating course assignment mechanisms as well some educational institutions focus on rank; e.g., Harvard Business School (HBS) emphasizes the average over the ranks of the ten courses that a student allocates on her reported rank-order list. Budish and Cantillon (2012) consider average rank, as well as other rank-based notions, to compare different course assignment mechanisms with the one used at HBS.

The literature on matching markets has thus considered rank as a natural summary statistic for quality of a match. A closely related notion that we will focus on is *average rank*, loosely speaking defined as the average, over all agents, of the rank of the object assigned to each agent on her list.[3] Average rank has been used, e.g., to compare stable assignments in two-sided markets where every agent ranks the agents on the other side independently and uniformly at random (Pittel, 1989; Ashlagi et al., 2017) or to evaluate core allocations in one-sided markets where agents rank objects uniformly at random (Knuth, 1996).

We similarly consider markets where agents on one side rank objects on the other side independently and uniformly at random. Our main finding is that, in such markets, there exists an assignment of objects to agents with a constant average rank. That is, there exists a constant R, independent of the market size, such that the object assigned to each agent has an average position of at most R on her list (Theorem 3.1). In addition, there exists a Bayesian Incentive Compatible mechanism that implements such an assignment (Section 5).

The proof of our main finding builds on techniques from random graph theory and probabilistic analysis. We use insights from a theorem of Walkup (1980) on the existence of perfect matchings in random graphs, together with two applications of the FKG inequality (Fortuin

[1] Teach For America assigns college graduates to teach in low-performing public and charter schools in the United States.

[2] This program places U.S. medical school students into residency training programs through a central mechanism that elicits rank-order lists from them. There, *match success* for an applicant is defined as being matched to the specialty of her first-ranked program (NRMP, 2020).

[3] The formal definition that we use is in Section 3.

et al., 1971), a fundamental correlation inequality in statistical mechanics and probabilistic combinatorics.[4] This inequality is used to keep track of the correlations between the objects' ranks on an agent's preference list.

Knuth (1996) is among the first to analyze the average rank in the market that we consider. He computes the exact value of the average rank under the *Random Serial Dictatorship* mechanism (RSD), in which agents are ordered randomly and choose objects one by one in that order (Abdulkadirouglu and Sönmez, 1998). His result implies that, in markets with an equal number of objects and agents, the average rank under RSD grows logarithmically in the market size.

Our result, together with that of Knuth (1996), indicates that the average rank can take a heavy toll under the RSD mechanism compared to the first-best. We further show that the gap between the first-best average rank and the average rank under RSD can persist when objects have capacities larger than one.[5,6] These findings highlight the possibility of using assignment methods other than RSD in scenarios where the average rank is a relevant objective. This is also aligned with the findings of Budish and Cantillon (2012) who, using field data from HBS, show that switching from the "draft" mechanism used at HBS to RSD increases the average rank of the ten courses in a student's schedule by a "large magnitude" (from 7.99 to 8.74).[7]

In Section 5 we consider the scenario where the agents' preference orders over objects are their private information, and we ask whether there exists a Dominant Strategy Incentive Compatible (DSIC) mechanism that implements the assignment with the first-best average rank. While the answer to this question turns out to be negative, we show that the answer is positive if the notion of DSIC is relaxed to Bayesian Incentive Compatibility (BIC). Together with Theorem 3.1, this implies the existence of BIC mechanisms in our main setup that implement assignments with a constant average rank of at most R.

Finally, we highlight the work of Che and Tercieux (2018) on the payoff equivalence of Pareto efficient assignments[8] and its connection to our findings. They show that when the agents' values for objects are independently and identically drawn (iid) from a distribution

[4] Section 3 recalls this inequality.

[5] That is when a bounded number of agents may be matched to the same object.

[6] More precisely, we prove that the first-best average rank does not grow infinitely large as the number of agents approaches infinity, whereas the average rank under RSD does, as long as objects' capacities grow sublogarithmically in the number of agents.

[7] Remarkably, while the draft mechanism is not strategyproof, they compute the average rank according to the students' *true* preferences, taken from a survey data conducted by the HBS administration.

[8] An assignment is Pareto efficient if there is no other assignment in which every agent is weakly better off with at least one agent being strictly better off.

with bounded support, all Pareto efficient assignments achieve the utilitarian upper bound in the large market limit, and thus are payoff equivalent.[9,10] At first glance this result may seem contradictory to the large gap between the average ranks of the RSD and the first-best assignments, as both assignments are Pareto efficient. However, the payoff equivalence result does not (need to) hold in our setup because the conditions required for it—the bounded support and the iid assumptions—do not hold here.[11]

2 Related literature

In the literature on matching markets, rank has been used as a standard notion of evaluating assignments. Knuth (1996) is one of the earliest to consider the notion of average rank in one-sided markets that are drawn uniformly at random. He analyzes the assignment model of Shapley and Scarf (1974) and provides an exact analysis for the average rank and its variance in that model. He also shows that these statistics are respectively equal to the average rank and its variance under the RSD mechanism.

The existence of assignments with a constant average rank in uniform markets has been shown in the literature when there is a surplus of objects that grows linearly in the market size. For instance, this a direct corollary of the work of Frieze and Melsted (2009) and Fountoulakis and Panagiotou (2012) on cuckoo hashing. Furthermore, it has been noted in the literature that in markets with a linear surplus of objects, the average rank of the assignment generated by RSD is constant.[12] Our results complement these findings by showing that the existence of assignments with a constant average rank does not depend on the level of market imbalance.

Featherstone (2020) defines an assignment to be rank efficient if its rank distribution cannot be stochastically dominated. He considers rank efficiency as a common objective of policy makers and shows that it can be attained by looking for local improvements that increase a natural objective. He also shows that no rank-efficient mechanism is strategyproof in general.[13]

[9]Their result is in fact more general: they show that the payoff equivalence result holds even in the presence of common-value components for objects, as long as the idiosyncratic components are drawn independently and uniformly at random from distributions with bounded support.

[10]Lee and Yariv (2018) establish a similar result in two-sided markets for stable assignments.

[11]The bounded support assumption does not hold because the rank of an object does not have a bounded range as the market grows large. The iid assumption does not hold because, if an object is ranked k-th on an agent's list, no other object can be ranked k-th.

[12]E.g., see Footnote 13 in Ashlagi et al. (2017)

[13]We complement this finding in Section 5 by considering all market sizes and object capacities.

The literature on *online assignment* problems has also considered rank as a notion of efficiency. Bruss (2000) develops optimal policies for the *hiring problem* (also known as the *secretary problem*) in which an employer needs to hire the top-ranked employee from a finite number of applicants, arriving in random order one by one. Chun and Sumichrast (2006) extend this setting to a matching setting for sequential assignment of jobs to machines.

In two-sided markets as well, the notions of average rank and rank distribution have been considered to evaluate assignment efficiency. Abdulkadirouglu et al. (2005) use data from the NYC high school match to empirically compare different tie-breaking rules for the Deferred Acceptance algorithm according to their rank distributions. Ashlagi et al. (2019) and Ashlagi and Nikzad (2020) compare the effect of tie-breaking rules on the students' rank distributions. Pittel (1992) and Ashlagi et al. (2017) analyze the average rank of agents in stable assignments in the context of marriage markets and characterize the stark effect of market imbalance on the average rank.[14]

Finally, we discuss relevant results in random graph theory and random matrix theory. From a technical perspective, we build on techniques from random graph theory and probabilistic analysis. Our proof uses insights from a theorem of Walkup (1980) on the existence of perfect matchings in random graphs, together with two applications of the FKG inequality (Fortuin et al., 1971), a fundamental correlation inequality in statistical mechanics and probabilistic combinatorics.

We also highlight the Parisi conjecture in random matrix theory, proved by Linusson and Wästlund (2004) and Nair et al. (2005) independently. The Parisi conjecture states that in an $n \times n$ matrix whose elements are iid from the exponential distribution with mean 1, there exists a set of n elements, no two in the same row or column, that asymptotically sum up to $\pi^2/6$. While this is reminiscent of Theorem 3.1 in some aspects, there is no formal connection to the extent of our knowledge.[15] Theorem 3.1 is concerned with the average rank, whereas the Parisi conjecture is concerned with the sum of iid real numbers. In particular, in the setting of the Parisi conjecture, the elements of the matrix are independent, whereas in our setting, the elements in a row (ranks of the objects) are correlated. For handling these correlations, our proof takes a completely different approach using the FKG inequality (Fortuin et al., 1971). Our approach also provides an upper bound—in the sense of first-order stochastic dominance—on the distribution of the rank of the object assigned to each

[14]Their results on the effect of market imbalance also extend to the number of stable partners.

[15]In Section ii of the online appendix, we discuss two natural but unsuccessful proof approaches based on the Parisi conjecture and the complexities that these approaches encounter. Making a formal connection to the Parisi conjecture remains an interesting theoretical direction.

agent (Remark 4.11).

3 Setup

A *matching market* contains a set of agents A and a set of objects O. We use n, m to denote $|A|, |O|$, respectively. Each object has a capacity c. Throughout, we assume that $n = cm$, unless explicitly stated otherwise. (We will see that the main theorem also holds without the latter assumption). An *assignment* is a function $\mu : A \to O$ that assigns each agent to an object without assigning more than c agents to any object.

Each agent has a strict[16] preference order over the objects, which is also called the *preference list* of the agent. The position of an object o on the preference list of an agent a is called the *rank* of the object for that agent, and it is denoted by $r_a(o)$. The *average rank* of an assignment μ is defined as

$$\overline{r}(\mu) = \frac{1}{|A|} \cdot \sum_{a \in A} r_a(\mu(a)).$$

The set of preference lists of all agents is called a *preference profile*. When A, O are known from the context, we denote the set of all preference profiles by Π and a typical member of it by π. For a preference profile π, let the *rank-optimal* assignment, $r^\star(\pi)$, be the assignment with the minimum average rank.

A *uniform market* is a matching market in which the preference order of each agent is drawn independently and uniformly at random from the set of all strict orderings of objects.

We next state some preliminaries required for the main findings.

Asymptotic notions We say a sequence of events $E_1, E_2, \ldots$ occurs *with high probability* (whp) if $\mathbb{P}[E_k]$ approaches 1 as k approaches infinity.

The FKG Inequality The Fortuin-Kasteleyn-Ginibre (FKG) inequality (Fortuin et al., 1971) is a correlation inequality. Informally, it says that an "increasing event" and a "decreasing event" are negatively correlated, while two "increasing events" are positively correlated.

Let $\mathcal{L}$ be a finite distributive lattice, $V(\mathcal{L})$ be the set of its elements, and $\mu : V(\mathcal{L}) \to \mathbb{R}_+$

[16] Strictness simplifies notation, but it is not a necessary assumption.

be a non-negative function that satisfies log-supermodularity, i.e., for any two $x, y \in V(\mathcal{L})$,

$$\mu(x \wedge y)\mu(x \vee y) \geq \mu(x)\mu(y),$$

where $\wedge, \vee$ are the meet and join operators of the lattice, respectively.

By the FKG inequality, when $f, g : V(\mathcal{L}) \to \mathbb{R}_+$ are respectively increasing and decreasing functions on the lattice $\mathcal{L}$, it holds that

$$\left(\sum_{x \in V(\mathcal{L})} f(x)g(x)\mu(x)\right)\left(\sum_{x \in V(\mathcal{L})} \mu(x)\right) \leq \left(\sum_{x \in V(\mathcal{L})} f(x)\mu(x)\right)\left(\sum_{x \in V(\mathcal{L})} g(x)\mu(x)\right).$$

The direction of the inequality is reversed when both functions are increasing (or decreasing).

3.1 Average Rank of Rank-Optimal Assignments

We next show that, in a uniform market, the expected average rank of the rank-optimal assignment is bounded from above by a constant independent of the market size.

Theorem 3.1. *There exists a constant R independent of c, n such that the expected average rank of the rank-optimal assignment in a uniform market is at most R.*

We discuss this result and the proof approach in Section 3.2. There, using numerical simulations we illustrate that the constant R is small (less than 2) for any capacity parameter c. The full proof of the theorem is presented in Section 4.

From this theorem it also follows that rank-optimal assignments have a constant expected average rank in uniform markets where there is a surplus or shortage of objects. (In case of a shortage of objects, the average rank is defined over the agents who are assigned an object.) This is formally proved in the online appendix, Section i.

3.2 Discussion of the Result and the Proof Approach

To prove Theorem 3.1, we show that the expected average rank of the rank-optimal assignment is smaller in a uniform market with n agents and n/c objects each with capacity c than in a uniform market with n agents and n objects each with capacity 1 (Lemma A.1). Thus, by this lemma, any constant R that bounds the expected average rank in markets with unit capacities will also be a valid upper bound in markets with larger capacities. Our proof shows that $R < 7\frac{3}{4}$ for sufficiently large n, whereas our simulations demonstrate that $R < 2$ (Figure 1).

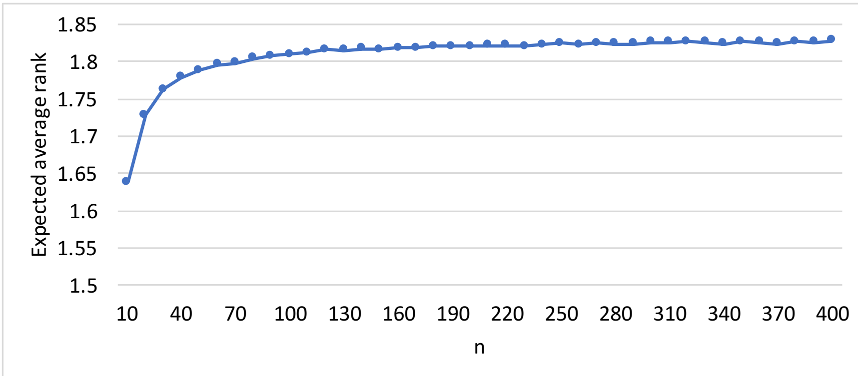

Figure 1: For each $n \in \{10, 20, 30, \ldots, 400\}$, we report the average over 1000 independent uniform markets with unit capacities. The largest reported average is below 1.83. In addition, for $n = 1000$ and $n = 10000$ we take the average over 10 independent draws due to computational limitations; the simulations report average ranks 1.831 and 1.827, respectively.

A natural proof approach for the case of unit capacities would be constructing a bipartite graph with agents on one side and objects on the other side, where every agent is connected to her k top-ranked objects, for some constant $k \geq 1$. The existence of a perfect matching[17] in such a graph would prove the claim. It turns out, however, for this graph to contain a perfect matching, k must grow logarithmically with the number of agents (Frieze and Karonski, 2012). This construction thus can only prove the existence of an assignment with an average rank of the order of $\log n$.[18,19]

Our proof starts with a different graph, which is inspired from a result of Walkup (1980). He shows that in a random bipartite graph with n nodes on each side, where each node is connected to two distinct neighbors independently and uniformly chosen from the other side, there exists a perfect matching with high probability. Inspired by this result, we construct a bipartite graph G with partitions A and O. In this graph, connect each node $a \in A$ to the three top-ranked objects on the preference list of a. We call these objects the *right-neighbors* of a. Also, connect every object o to every agent who ranks o worse than 3 and no worse

[17] We recall that a matching in a graph is a set of edges with no two edges having a node in common. A perfect matching is one that covers every node of the graph.

[18] Remarkably, the average rank generated by RSD is also of the same order.

[19] When there is a surplus of objects that grows linearly in n, Fountoulakis and Panagiotou (2012) show that there exists a constant k such that the constructed graph contains a perfect matching with high probability.

than r_o, where r_o is defined as follows. For object $o \in O$, let r_o be the smallest integer larger than 3 such that there exists at least 3 agents for whom o has a rank (strictly) worse than 3 and no worse than r_o. If there is no such integer, then define $r_o = n$. The agents who rank o worse than 3 and no worse than r_o are called the *left-neighbors* of o.[20]

Intuitively, the graph that we construct G has similarities with the graph considered in Walkup (1980), in the sense that each agent is adjacent to three distinct right-neighbors chosen *independently* and *uniformly at random*. Unlike in Walkup (1980), however, the left-neighbors of objects are not chosen independently: the objects' left-neighbors may not be independent from each other or from the agents' right-neighbors, as we elaborate later.

In what follows we describe how G is used to prove the theorem. The proof has two steps. In step A, we show that G has a perfect matching with high probability, and in step B, we show that the perfect matching has a constant average rank.

Step A This step shows that G has a perfect matching with probability at least $1 - n^{-3}$ for all $n \geq 50$. As in Walkup (1980), the proof approach is based on the König theorem. By this theorem, the size of the maximum matching plus the size of the maximum independent set in a bipartite graph equals the total number of nodes.[21] Therefore, to prove that G has a matching of size n, it suffices to show that it has no independent set of size $n + 1$. We prove that the probability that G contains an independent set of size $n + 1$ is at most n^{-3}. This is done by bounding the probability that a fixed subset of nodes of size $n + 1$ forms an independent set, and then taking a union bound over all subsets of nodes of size $n + 1$.

Bounding the probability that a fixed subset of nodes of size $n + 1$ forms an independent set is relatively straightforward when the neighbors of a node are chosen independently, as in Walkup (1980). In the graph that we construct, however, the neighbors of nodes are *not* chosen independently. In particular, there are two types of correlations involved:

i. *Correlations across agents and objects.* The left-neighbors of a node $o \in O$ are not independent of the right-neighbors of a node $a \in A$.
ii. *Correlations between objects.* The left-neighbors of a node o_1 are not independent of the left-neighbors of a node o_2.

To bound the probability that a fixed subset of nodes of size $n + 1$ forms an independent

[20] In the bipartite graph that Walkup (1980) studies, each node is connected to d other nodes drawn independently and uniformly at random from the other side. This graph cannot be used to prove our main theorem, because every "left-neighbor" of an object in the graph would rank that object $\frac{m+1}{2}$-th on her list, on average. Thus, a perfect matching in this graph would not necessarily have a constant average rank.

[21] We recall that the size of a matching equals its number of edges and that an independent set in a graph is a subset of nodes that are pairwise non-adjacent.

set, we need to handle the above correlations. This is the bulk of the analysis in this step, and involves applications of the FKG inequality, as outlined next.

Suppose $X \subseteq A$ and $Y \subseteq O$ are aribtrary subsets of nodes of G such that $|X|+|Y| = n+1$. We will provide an upper bound on $p_{(X,Y)}$, the probability that $X \cup Y$ forms an independent set in G. Let R denote the event that there is no node in X which has a right-neighbor in Y. Also, let L denote the event that there is no node in Y which has a left-neighbor in X. Thus, $p_{(X,Y)} = \mathbb{P}[L \cap R]$. We will show that $\mathbb{P}[L \cap R] \leq \mathbb{P}[L]\,\mathbb{P}[R]$, and then will bound each of the terms on the right-hand side separately.

The inequality $\mathbb{P}[L \cap R] \leq \mathbb{P}[L]\,\mathbb{P}[R]$ holds intuitively because, by the construction of the bipartite graph, the event that an object o is not a right-neighbor of an agent a is negatively correlated with the event that the agent a is not a left-neighbor of the object o. We prove this negative correlation inequality formally using the FKG inequality (Claim 4.2). It remains to provide upper bounds on $\mathbb{P}[L]$ and $\mathbb{P}[R]$.

We can derive a closed-form expression for $\mathbb{P}[R]$, essentially because the right-neighbors of every node $a \in X$ are chosen independently of the right-neighbors of every other node $a' \in X$. To provide an upper bound on $\mathbb{P}[L]$, let L_y denote the event that an object $y \in Y$ has no left-neighbor in X. We first show that $\mathbb{P}[L] \leq \prod_{y \in Y} \mathbb{P}[L_y]$ (Claim 4.5). This holds intuitively because, by the construction of the bipartite graph, the event that an agent is a left-neighbor of the object y is negatively correlated with the event that the agent is a left-neighbor of another object $y' \neq y$. Finally, we provide an upper bound on $\mathbb{P}[L_y]$, which gives an upper bound on $\mathbb{P}[L]$, and thereby the promised upper bound on $p_{(X,Y)}$ (Claim 4.8).

Step B This step is relatively simpler. We show that a perfect matching in G has an expected average rank of at most $7\frac{3}{4}$ when n is sufficiently large. Define the *weight* of an edge (a, o) in G, denoted by $w_{a,o}$, to be the rank of object o on agent a's preference list. To prove the theorem, we show that the expected rank of the maximum-weight edge adjacent to an object is bounded by $7\frac{3}{4}$. To this end, we first note that $w_{a,o} \leq 3$ if o is a right-neighbor of a and $w_{a,o} \geq 4$ if a is a left-neighbor of o. Thus, it suffices to show that

$$\mathbb{E}\left[\max_{a \in A}\{w_{a,o} : a \text{ is a left-neighbor of } o\} \;\middle|\; \text{o has a left-neighbor}\right] \leq 7\frac{3}{4}.$$

This bound is proved in Claim 4.10. To provide intuition for it, we consider the case where o has exactly 3 left-neighbors. These are the 3 agents who, among all agents that rank o worse than 3, rank o the highest. Let these 3 agents be a_1, a_2, a_3, and let r_i denote the rank of o on agent a_i's list. Without loss of generality, suppose $r_1 \leq r_2 \leq r_3$. The above bound then boils down to $\mathbb{E}[r_3] \leq 7\frac{3}{4}$, the proof for which is sketched below.

Since the market is uniform, the chance that no agent ranks o fourth is at most $(1-\frac{1}{n})^n$; the chance that no agent ranks o fourth or fifth is at most $(1-\frac{1}{n})^{2n}$; and similarly, the chance that no agent ranks o fourth, fifth,..., k-th is at most $(1-\frac{1}{n})^{(k-3)n}$. Therefore, $r_1 - 3$ has a distribution that is stochastically dominated by a geometric random variable with failure probability $q_1 = (1-\frac{1}{n})^n$. A similar argument shows that $r_2 - r_1$ and $r_3 - r_2$ respectively have distributions that are stochastically dominated by geometric random variables with failure probabilities $q_2 = (1-\frac{1}{n})^{n-1}$ and $q_3 = (1-\frac{1}{n})^{n-2}$. Therefore,

$$\mathbb{E}[r_3] \leq 3 + \frac{1}{1-q_1} + \frac{1}{1-q_2} + \frac{1}{1-q_3}.$$

This concludes the proof since the right-hand side approaches $3(1+\frac{1}{1-1/e}) \approx 7.745$ as n approaches infinity.[22]

4 Proof of Theorem 3.1

We first define the preliminary notions. Some of the notions that we use in the following analysis were already defined in the proof sketch (Section 3.2); however, we include all such definitions below to make this section independently readable.

For every finite set S, let $U(S)$ denote the uniform distribution over S. Also, let $\Delta(S)$ denote the probability simplex defined over S, i.e., the set of all probability distributions defined over S. For every positive integer i, let $[i]$ denote the set $\{1, \dots, i\}$.

We recall that the set of all possible preference profiles in a market is denoted by Π. The preference profile π of a uniform market is drawn from $U(\Pi)$. Thus, unless stated otherwise, it is assumed in the following analysis that the probability space is defined by the sample space Π, the event space 2^Π, and the probability function that assigns probability $\frac{|\Psi|}{|\Pi|}$ to an event $\Psi \subseteq \Pi$.

We next prove Theorem 3.1 for the case of unit capacities ($c = 1$). The proof for non-unit capacities ($c > 1$) is a corollary of the proof for unit capacities and is relegated to the appendix, Section A.1.

Throughout the proof, let $d = 3$. First, from a given preference profile, we construct a bipartite graph G with A, O being the set of nodes on each side. We call A the left side of

[22] The convergence to the limit is from above.

the graph and O the right side. The set of edges in the graph is $E = \overleftarrow{E} \cup \overrightarrow{E}$, where $\overleftarrow{E}$ and $\overrightarrow{E}$ are defined next. $\overrightarrow{E}$ (the set of edges from left to right) contains an edge (a, o) for every agent a and every object o that is ranked d or better[23] on a's preference list.

We define $\overleftarrow{E}$ (the set of edges from right to left) as follows. For every object $o \in O$, define r_o to be the smallest integer larger than d such that there exists at least d agents for whom o has a rank (strictly) worse than d and no worse than r_o. If there is no such integer, then define $r_o = n$. For every object o, $\overleftarrow{E}$ contains an edge (o, a) for every agent a that ranks o worse than d and no worse than r_o. (Figure 2 illustrates this definition.)

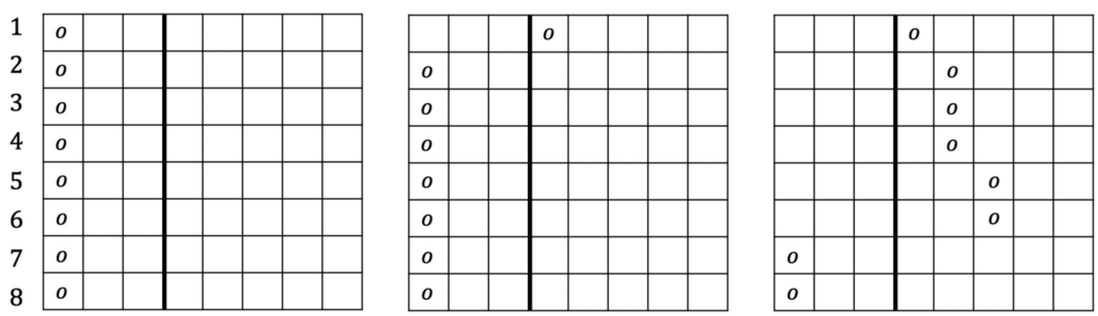

Figure 2: $n = 8$ and $d = 3$. Each row corresponds to an agents' preference list. In the left panel, object o is ranked first by all agents; thus, $\overleftarrow{E}$ would contain no edge adjacent to o. In the middle panel, o is ranked worse than third only by agent 1; thus, $\overleftarrow{E}$ would contain a single edge adjacent to o, $(1, o)$. In the right panel, 6 (i.e., at least d) agents rank o worse than third; $r_o = 5$, and thus, $\overleftarrow{E}$ would contain 4 edges adjacent to o, namely $(1, o), (2, o), (3, o), (4, o)$.

The proof is done in two steps. In Step A we show that G contains a perfect matching whp, and then we use this fact in Step B to provide an upper bound on the expected average rank of the rank-optimal matching.

4.1 Step A: Existence of a Perfect Matching with High Probability

A k-tuple is a pair (X, Y) with $X \subseteq A$ and $Y \subseteq O$ such that $|X| = k$ and $|Y| = n - k + 1$. A k-tuple (X, Y) is *independent* if $(X \times Y) \cap E = \varnothing$. We recall that, by König's theorem,[24] G contains a matching of size n if and only if it does not contain an independent k-tuple for any positive integer $k \leq n$. Let p_k denote the probability that G contains an independent k-tuple, and define $p(n) = \sum_{k=1}^{n} p_k$. To complete this step, we will prove that $p(n) \leq n^{-3}$

[23] A rank d or better is a rank that belongs to the set $\{1, \ldots, d\}$.

[24] See, e.g., West (2000).

holds for $n \geq 50$. We next present the proof outline and then the complete proof for this inequality.

By $p_{(X,Y)}$ denote the chance that a k-tuple (X, Y) is independent. Let L, R respectively denote the events that $(X \times Y) \cap \overset{\curvearrowleft}{E} = \varnothing$ and $(X \times Y) \cap \overset{\curvearrowright}{E} = \varnothing$. (We recall that every event is, by definition, a subset of Π; therefore, we treat L, R as such subsets.) Define L_i to be the event that there is no edge in $\overset{\curvearrowleft}{E}$ that connects $i \in Y$ to an element of X.

We provide an upper bound on $p_{(X,Y)}$, and then take a union bound over all k-tuples to bound p_k. This is done as follows. Observe that, $p_{(X,Y)} = \mathbb{P}\left[L \cap R\right] = \mathbb{P}\left[R\right]\mathbb{P}\left[L|R\right]$. We first write a closed-form expression for $\mathbb{P}\left[R\right]$ in Claim 4.1. Then, we prove the negative correlation inequalities $\mathbb{P}\left[L|R\right] \leq \mathbb{P}\left[L\right]$ and $\mathbb{P}\left[L\right] \leq \prod_{i \in Y} \mathbb{P}\left[L_i\right]$ in Claim 4.2 and Claim 4.5, respectively. This would imply that $p_{(X,Y)} \leq \mathbb{P}\left[R\right] \prod_{i \in Y} \mathbb{P}\left[L_i\right]$. Providing a closed-form upper bound on $\mathbb{P}\left[L_i\right]$ (Claim 4.8) then gives the promised upper bound on $p_{(X,Y)}$. Then, taking a union bound over all k-tuples (X, Y) gives an upper bound on p_k, which we use to show that $\sum_{k=1}^{n} p_k \leq n^{-3}$ (Fact 4.9).

For the following claims, without loss of generality, we suppose that $X = \{1, \ldots, k\}$ and $Y = \{1, \ldots, l\}$, where $l = n - k + 1$.

Claim 4.1. *For a k-tuple (X, Y),*

$$p_{(X,Y)} = \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot \mathbb{P}\left[L|R\right]. \tag{4.1}$$

Proof. We first note that $p_{(X,Y)} = \mathbb{P}\left[L \cap R\right] = \mathbb{P}\left[R\right]\mathbb{P}\left[L|R\right]$. Thus, to prove the claim it suffices to show that $\mathbb{P}\left[R\right] = \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k$. To this end, for every agent $x \in X$, define R_x as the event that agent x lists no object in Y as one of her top d objects. Thus, $R = \cup_{x \in X} R_x$. Since the preference list of each agent is distributed uniformly over the set of all strict orderings of objects, then

$$\mathbb{P}\left[R_x\right] = \frac{\binom{n-|Y|}{d}}{\binom{n}{d}} = \frac{\binom{k-1}{d}}{\binom{n}{d}}.$$

To complete the proof, we observe that since the agents' preference lists are iid, then

$$\mathbb{P}\left[R\right] = \prod_{x \in X} \mathbb{P}\left[R_x\right] = \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k.$$

$\square$

Claim 4.2 (Negative correlation of L, R)**.** $\mathbb{P}[L|R] \leq \mathbb{P}[L]$.

Proof. We will use the FKG inequality to prove the claim. To define the lattice required for applying the FKG inequality, we first define the set of its elements, V_l. Each element $x \in V_l$ is a sequence $(x^1, \ldots, x^n)$, where x^i is a subset of $[n]$ with size l. (See Figure 3.)

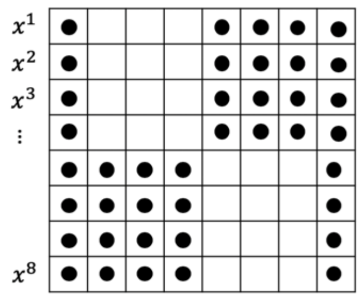

Figure 3: For $n = 8$ and $l = 5$, we illustrate an element $x = (x^1, \ldots, x^n)$ of the lattice where $x^i = \{1, 5, 6, 7, 8\}$ for $i \leq 4$ and $x^i = \{1, 2, 3, 4, 8\}$ for $i \geq 5$. The columns from left to right respectively correspond to objects $1, \ldots, n$.

To define the lattice order over V_l, we first need a few definitions. For any object i, let $h(i) = i$ if $i > d$ and let $h(i) = n + i$ otherwise. Define the total order $\preceq$ on the set of objects $O = \{1, \ldots, n\}$ by $i \preceq j$ if $h(i) \leq h(j)$. With slight abuse of notation, for any two vectors of the same size, namely $u = (u_1, \ldots, u_s), v = (v_1, \ldots, v_s)$, we write $u \preceq v$ if $u_j \preceq v_j$ holds for all j. By $\overrightarrow{u}$ we denote vector that is obtained by sorting the elements of u in an increasing order with respect to $\preceq$.

For any $x, y \in V_l$, we define $x \preceq_l y$ when

$$\begin{cases} \overrightarrow{y^a} \preceq \overrightarrow{x^a}, \ \forall a \in X \\ \overrightarrow{x^a} \preceq \overrightarrow{y^a}, \ \forall a \in A \backslash X. \end{cases} \tag{4.2}$$

The next fact shows that $\preceq_l$ indeed defines a lattice over V_l. Figure 4 provides a graphical representation of this lattice.

Fact 4.3. *The relation $\preceq_l$ is a lattice over V_l.*

This fact is proved in the appendix, Section A. The intuition is that $\preceq_l$ is the Cartesian product of two lattices: one that orders tuples $(x^1, \ldots, x^k)$ and one that orders tuples $(x^{k+1}, \ldots, x^n)$.

We denote by $\mathcal{L}[l]$ the lattice defined by $\preceq_l$. Since l is fixed in the proof of this claim, we drop the argument and denote the lattice by $\mathcal{L}$ throughout the proof of this claim.

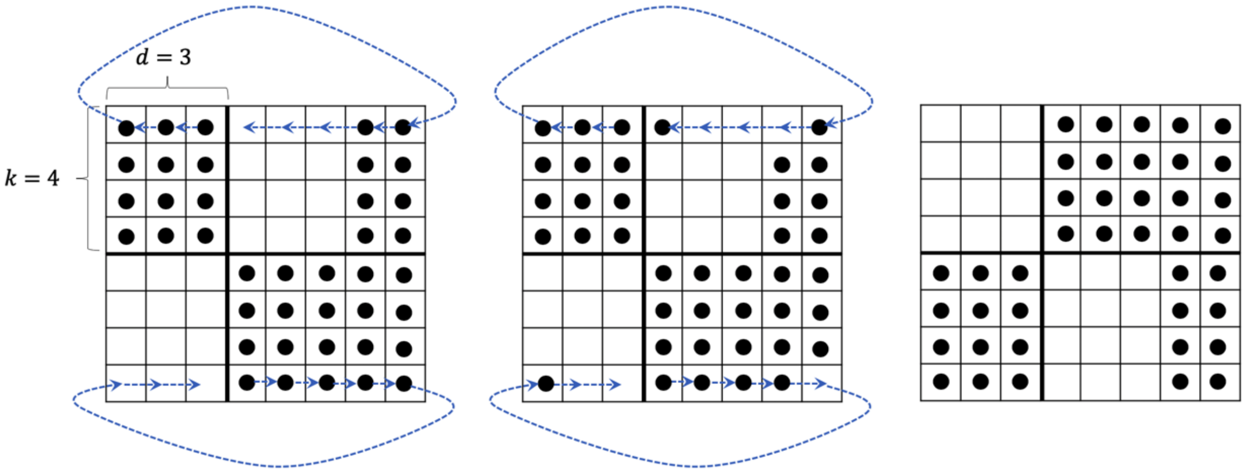

Figure 4: $n = 8$ and $l = 5$. Panels from left to right: the smallest element of the lattice, an element between the smallest and the largest element, and the largest element. The elements are represented in the same fashion as in Figure 3. Consider each black circle to be a *chip*. One can *move* from a smaller to a larger element in the lattice by moving a chip to an empty cell in the direction of the arrows. Since $k = 4$, the direction is the same for the first four rows, and is reversed for the last four rows.

We associate each preference profile π with an element of the lattice, which we denote by $\mathcal{L}_\pi = (\mathcal{L}_\pi^1, \ldots, \mathcal{L}_\pi^n)$, where $\mathcal{L}_\pi^a$ is the set of ranks of the objects $1, \ldots, l$ on agent a's preference list. For every element x of the lattice, define

$$\Pi_x = \{\pi \in \Pi : x = \mathcal{L}_\pi\}.$$

Define the functions $f_L, f_R : \mathcal{L} \to [0, 1]$

$$f_L(x) = \mathbb{P}_{\pi \sim U(\Pi_x)}\left[\pi \in L\right],$$
$$f_R(x) = \mathbb{P}_{\pi \sim U(\Pi_x)}\left[\pi \in R\right].$$

In words, $f_L(x)$ is the probability that event L occurs at a preference profile π drawn uniformly at random from the set of all preference profiles whose associated element of the lattice is x. Similarly, $f_R(x)$ gives that probability but for the event R instead of L.

Fact 4.4. *For every $x \in V_l$, $f_R(x) \in \{0, 1\}$; also, for every $y \in V_l$ with $x \preceq_l y$, $f_R(x) \leq f_R(y)$ and $f_L(x) \geq f_L(y)$.*

The proof for this fact is in the appendix, Section A. The intuition for the first part is that, for any preference profile π, its associated lattice element $\mathcal{L}_\pi$ has all the information needed (i.e., the ranks of objects in Y on agents' lists) to evaluate whether $\pi \in R$. The proof for the second part is technical. A loose intuition can be given through Figure 4: as one transforms x to y by moving the "chips" one by one along the direction of the arrows, (i) the total number of preference profiles in Π_x remains the same (i.e., $|\Pi_x| = |\Pi_y|$), (ii) the number of preference profiles in Π_x that are a member of R increases, because moving a chip along the direction of the arrows may eliminate existing edges in $\widehat{E}$ but does not create any new ones, and (iii) the number of preference profiles that are a member of L decreases, because by moving a chip along the direction of the arrows one cannot transform the set $\widehat{E} \cap (X \times Y)$ from being nonempty into being empty.

Given Fact 4.4, we are now ready to apply the FKG inequality on the lattice $\mathcal{L}$ with the functions f_L and f_R and the uniform probability distribution $U(V_l)$ defined over the lattice elements. By Fact 4.4, the FKG inequality implies the following *negative correlation* bound:

$$\mathbb{E}_{x\sim U(V_l)}\left[f_L(x)f_R(x)\right] \leq \mathbb{E}_{x\sim U(V_l)}\left[f_L(x)\right]\mathbb{E}_{x\sim U(V_l)}\left[f_R(x)\right].$$

We next observe that when $\pi \sim U(\Pi)$, then the lattice element $\mathcal{L}_\pi$ is distributed uniformly at random over V_l, by symmetry. Therefore, we can rewrite the above inequality as

$$\mathbb{E}_{\pi\sim U(\Pi)}\left[f_L(\mathcal{L}_\pi)f_R(\mathcal{L}_\pi)\right] \leq \mathbb{E}_{\pi\sim U(\Pi)}\left[f_L(\mathcal{L}_\pi)\right]\mathbb{E}_{\pi\sim U(\Pi)}\left[f_R(\mathcal{L}_\pi)\right]. \tag{4.3}$$

Since $f_R(\mathcal{L}_\pi) \in \{0,1\}$ holds for all π by Fact 4.4, then the left-hand side of the above inequality equals $\mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L \cap R\right]$. On the right-hand side, we note that

$$\mathbb{E}_{\pi\sim U(\Pi)}\left[f_L(\mathcal{L}_\pi)\right] = \mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L\right], \quad \mathbb{E}_{\pi\sim U(\Pi)}\left[f_R(\mathcal{L}_\pi)\right] = \mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in R\right]$$

hold by the definition of f_L, f_R. Therefore, we can rewrite (4.3) as

$$\mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L \cap R\right] \leq \mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L\right]\mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in R\right].$$

This means $\frac{\mathbb{P}_{\pi\sim U(\Pi)}[\pi\in L\cap R]}{\mathbb{P}_{\pi\sim U(\Pi)}[\pi\in R]} \leq \mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L\right]$, which is the promised claim. □

Recall that L_i is the event that there is no edge in $\widehat{E}$ that connects $i \in Y$ to an element of X.

Claim 4.5 (Negative correlation of $\{L_i\}_{i\in Y}$). $\mathbb{P}\left[L\right] \leq \prod_{i=1}^{l} \mathbb{P}\left[L_i\right]$.

Proof. The proof is by induction. For any $j \leq l$, we will show that

$$\mathbb{P}\left[L_1 \cap L_2 \cap \ldots \cap L_j\right] \leq \prod_{i=1}^{j} \mathbb{P}\left[L_i\right].$$

The induction basis for $j = 1$ is trivial. The induction step supposes that, for some $l' < l$, the claim holds for all $j \leq l'$ and then proves the claim for $j = l' + 1$.

For the induction step, we need the following definitions. Recall (4.2), which defines the lattice $\mathcal{L}[l]$, parameterized by the parameter l. The lattice that we will use in the induction step is $\mathcal{L}[l']$. Thus, $V_{l'}$ is the set of the lattice elements; an element of the lattice is $(x^1, \ldots, x^n)$, where x^i is a subset of $[n]$ with size l'; and the lattice elements are ordered according to (4.2). For notational simplicity, we denote the lattice $\mathcal{L}[l']$ by $\mathcal{K}$.

We associate a preference profile π with an element of the lattice, denoted by $\mathcal{K}_\pi = (\mathcal{K}_\pi^1, \ldots, \mathcal{K}_\pi^n)$, where $\mathcal{K}_\pi^a$ is the set of ranks of the objects $1, \ldots, l'$ on agent a's preference list. Let $\Pi_x = \{\pi \in \Pi : x = \mathcal{K}_\pi\}$.

The induction step. The induction step will apply the FKG inequality on the lattice $\mathcal{K}$ and the function $f, g : V(\mathcal{K}) \to \mathbb{R}_+$, defined as follows:

$$\begin{aligned} f(x) &= \mathbb{P}_{\pi \sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'}\right], \\ g(x) &= \mathbb{P}_{\pi \sim U(\Pi_x)}\left[\pi \in L_{l'+1}\right]. \end{aligned}$$

Fact 4.6. *For every $x, y \in V(\mathcal{K})$ with $x \preceq_{l'} y$, $f(x) \geq f(y)$ and $g(x) \leq g(y)$.*

The proof is in Section A of the appendix. The intuition for the first inequality is similar to the intuition for $f_L(x) \geq f_L(y)$ from Fact 4.4. A loose intuition for the second inequality can be given through Figure 4: as one transforms x to y by moving the "chips" one by one along the direction of the arrows, the total number of preference profiles in Π_x remains the same (i.e., $|\Pi_x| = |\Pi_y|$) but the number of preference profiles that are a member of $L_{l'+1}$ increases, because by moving a chip along the direction of the arrows one cannot transform the set $\overset{\frown}{E} \cap (X \times \{l' + 1\})$ from being empty into being nonempty.

By Fact 4.6, the FKG inequality implies the following negative correlation inequality:

$$\mathbb{E}_{x \sim U(V(\mathcal{K}))}\left[f(x) g(x)\right] \leq \mathbb{E}_{x \sim U(V(\mathcal{K}))}\left[f(x)\right] \mathbb{E}_{x \sim U(V(\mathcal{K}))}\left[g(x)\right]. \tag{4.4}$$

Fact 4.7. *For every $x \in V(\mathcal{K})$,*

$$\mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'+1}\right] = \mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'}\right] \cdot \mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_{l'+1}\right].$$

The above fact is proved in Section A of the appendix. The intuition is that conditioning on the agents' ranks for the object $l'+1$ in π does not affect the probability of the event $\pi \in L_1 \cap \ldots \cap L_{l'}$, since $\pi \sim U(\Pi_x)$ and $x \in V(\mathcal{K})$.

By the definition of f, g,

$$\begin{aligned}
&\mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L_1 \cap \ldots \cap L_{l'+1}\right] \\
=&\mathbb{E}_{x\sim U(V(\mathcal{K}))}\left[\mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'+1}\right]\right] \\
=&\mathbb{E}_{x\sim U(V(\mathcal{K}))}\left[\mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'}\right] \cdot \mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_{l'+1}\right]\right] \\
\leq&\mathbb{E}_{x\sim U(V(\mathcal{K}))}\left[\mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_1 \cap \ldots \cap L_{l'}\right]\right] \cdot \mathbb{E}_{x\sim U(V(\mathcal{K}))}\left[\mathbb{P}_{\pi\sim U(\Pi_x)}\left[\pi \in L_{l'+1}\right]\right] \\
=&\mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L_1 \cap \ldots \cap L_{l'}\right] \cdot \mathbb{P}_{\pi\sim U(\Pi)}\left[\pi \in L_{l'+1}\right],
\end{aligned}$$

where in the first equality we have used the fact that the distribution imposed on π when $x \sim U(V(\mathcal{K}))$ and $\pi \sim U(\Pi_x)$ is just the uniform distribution over Π. The second equality is by Fact 4.7, and the inequality is by (4.4). The above bound completes the induction step and proves the claim. □

The next claim gives a closed-form upper bound for $p_{(X,Y)}$ by bounding the right-hand side of (4.1) from above. We then use this claim to provide a closed-form upper bound on p_k, and prove that $\sum_{k=1}^{n} p_k \leq n^{-3}$.

Claim 4.8.

$$p_{(X,Y)} \leq \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k (\omega+\delta)^l, \tag{4.5}$$

where $\omega = \binom{n-k}{d}/\binom{n}{d}$ and $\delta = \frac{d^{n-2d+1}e^d}{n^{n-2d}}$.

Proof. By (4.1) and Claims 4.2 and 4.5, it suffices to show that $\mathbb{P}\left[L_i\right] \leq \omega + \delta$ for $i \in Y$. To this end, let D be the event in which object i has a rank worse than d in at least d of the agents' preference lists. Observe that

$$\mathbb{P}\left[L_i\right] = \mathbb{P}\left[L_i | D\right] \times \mathbb{P}\left[D\right] + \mathbb{P}\left[L_i | \overline{D}\right] \times \mathbb{P}\left[\overline{D}\right].$$

We will show that (i) $\mathbb{P}\left[\overline{D}\right] \leq \delta$ and (ii) $\mathbb{P}\left[L_i | D\right] \leq \omega$, which would imply that $\mathbb{P}\left[L_i\right] \leq \omega + \delta$.

(i) $\mathbb{P}\left[\overline{D}\right] \leq \delta$. Let $\overline{D}_j$ be the event in which object i is ranked worse than d by exactly j agents. Observe that

$$\mathbb{P}\left[\overline{D}_j\right] = \binom{n}{j}\left(\frac{d}{n}\right)^{n-j}\left(1 - \frac{d}{n}\right)^{j} \leq \binom{n}{j}\left(\frac{d}{n}\right)^{n-j}.$$

Therefore, we can write

$$\mathbb{P}\left[\overline{D}\right] = \sum_{j=0}^{d-1} \mathbb{P}\left[\overline{D}_j\right] \leq d\binom{n}{d}\left(\frac{d}{n}\right)^{n-d} \leq \frac{d^{n-2d+1}e^{d}}{n^{n-2d}} = \delta, \tag{4.6}$$

where the first inequality holds since $d = 3 \leq \frac{n}{2}$, and the second one holds since $\binom{n}{d} \leq \left(\frac{ne}{d}\right)^{d}$ (Das, 2016).

(ii) $\mathbb{P}\left[L_i | D\right] \leq \omega$. Let A_j be the set of agents who assign rank j to object i, and let $A^h = \cup_{j=d+1}^{h} A_j$ be such that h is the smallest integer for which $|A^h| \geq d$. Such h exists when the event D holds. Observe that the set A^h is a random variable whose distribution, conditional on its size being s, is the uniform distribution over the set of all subsets of A with size s. (This holds because the agents' preference lists are iid from the uniform distribution.) Therefore, for every integer $g \geq 1$, we can write

$$\mathbb{P}\left[L_i \middle| \, g = |A^h|\right] = \frac{\binom{n-k}{g}}{\binom{n}{g}} = \prod_{j=1}^{g} \frac{n-k-j+1}{n-j+1},$$

where in the middle term the numerator is the number of subsets of $A \backslash X$ of size g, and the denominator is the number of subsets of A of size g. Recall that event D holds if and only if h exists (and if h exists, then $|A^h| \geq d$ would also hold). Since the right-hand side of the above equation is decreasing in g, and since $|A^h| \geq d$, then

$$\mathbb{P}\left[L_i | D\right] \leq \frac{\binom{n-k}{d}}{\binom{n}{d}} = \omega,$$

which is the promised bound. □

We now use Claim 4.8 and a union bound over all k-tuples (X, Y) with $|X| = k$ to write

$$p_k \le \binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k (\omega+\delta)^l$$
$$\le \binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \left(\omega^l + 2^l\delta\right),$$

where recall that $l = n - k + 1$, and the last inequality holds since the term $(\omega+\delta)^l$ has 2^l summands when expanded and $\omega, \delta \le 1$. We therefore can bound $p(n)$ as follows:

$$p(n) = \sum_{k=1}^{n} p_k \le \sum_{k=1}^{n} \binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \left(\omega^l + 2^l\delta\right)$$
$$= \sum_{k=1}^{n} \binom{n}{k}\binom{n}{k-1}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot \omega^l + \sum_{k=1}^{n} \binom{n}{k}\binom{n}{k-1}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot 2^l\delta. \quad (4.7)$$

Let $S_1(n)$ and $S_2(n)$ respectively denote the first and the second summand in (4.7). Walkup (1980) shows that[25]

$$S_1(n) \le \frac{1}{122}\left(\frac{d}{n}\right)^{(d+1)(d-2)}. \quad (4.8)$$

We finish the proof by providing an upper bound on $S_2(n)$. Since $\binom{n}{j} \le 2^n$ for all positive integers j, and $\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \le 1$, then

$$S_2(n) \le 2^{3n}\delta = 2^{3n} \cdot \frac{d^{n-2d+1}e^d}{n^{n-2d}}. \quad (4.9)$$

The following fact uses (4.8) and (4.9) to bound $S_1(n) + S_2(n)$ from above.

Fact 4.9. *For $d = 3$, $S_1(n) + S_2(n) \le n^{-3}$ for $n \ge 50$.*

Proof. When $d = 3$ and $n \ge 50$, (4.8) implies that

$$S_1(n) \le \frac{81}{122n}\left(\frac{1}{n}\right)^3 \le \frac{1}{2n^3}. \quad (4.10)$$

[25]This is the last bound in the proof of Theorem 1 in Walkup (1980).

Also, (4.9) implies that

$$S_2(n) \leq 2^{3n} \cdot \frac{3^{n-5}e^3}{n^{n-6}}. \tag{4.11}$$

Let $g(n)$ denote the right-hand side of the above inequality. Define $h(n) = 2n^3 g(n)$. We observe that

$$h'(n) = e^3 2^{3n+1} 3^{n-5} n^{8-n}(n(\log(24) - 1) - n\log(n) + 9).$$

Since

$$n(\log(24) - 1) - n\log(n) + 9 < 0$$

for $n \geq 50$, then $h'(n) < 0$ for such n. Therefore, $h(n) \leq h(50) < 1$. Recall that $h(n) = 2n^3 g(n)$. Thus, $g(n) \leq \frac{1}{2n^3}$ for $n \geq 50$. Together with (4.11), this implies that $S_2(n) \leq \frac{1}{2n^3}$. This bound, together with (4.10), implies that $S_1(n) + S_2(n) \leq n^{-3}$ for $n \geq 50$. $\square$

By (4.7), $p(n) \leq S_1(n) + S_2(n)$. By Fact 4.9, the right-hand side is at most n^{-3} when $n \geq 50$. This proves the promised bound on $p(n)$ and completes step A.

4.2 Step B: Bounding the Average Rank in the Perfect Matching

Define the *weight* of an edge (a, o) in the graph G to be the rank of object o on agent a's preference list. We showed that G has a perfect matching with probability at least $1 - n^{-3}$ when $n \geq 50$. To bound the expected average rank in the matching, we provide an upper bound on $\sum_{o \in O} w_o$, where w_o is the weight of the maximum-weight edge adjacent to $o \in O$. (Define $w_o = 0$ if there are no edges adjacent to o.)

For any object o, we provide an upper bound on $\mathbb{E}[w_o]$, where the expectation is taken over $\pi \sim U(\Pi)$. Since all of the expectation and probability operators are with respect to $\pi \sim U(\Pi)$ in the rest of the proof, we drop this subscript from the notation.

Let $w = \sum_{o \in O} w_o$, and $\mathfrak{m}$ denote the event that G has a perfect matching. Since $\mathbb{P}[\mathfrak{m}]\,\mathbb{E}[w|\mathfrak{m}] + \mathbb{P}[\bar{\mathfrak{m}}]\,\mathbb{E}[w|\bar{\mathfrak{m}}] = \mathbb{E}[w]$, then $\mathbb{P}[\mathfrak{m}]\,\mathbb{E}[w|\mathfrak{m}] \leq \mathbb{E}[w]$. Because $\mathbb{P}[\mathfrak{m}] \geq 1 - n^{-3}$ as shown in Step A, then

$$\mathbb{E}[w|\mathfrak{m}] \leq \frac{\mathbb{E}[w]}{1 - n^{-3}}.$$

Therefore, to complete Step B, it suffices to bound $\mathbb{E}[w]$ from above.

The next claim will prove that $\mathbb{E}[w] \leq 3n(1 + \frac{1}{1-(1-\frac{1}{n})^{n-2}})$. This would complete the proof because this inequality together with the above bound implies that, for $n \geq 50$,

$$\frac{1}{n}\mathbb{E}\left[w|\mathfrak{m}\right] \leq \frac{3(1+\frac{1}{1-(1-\frac{1}{n})^{n-2}})}{1-n^{-3}} < 7.9,$$

where the last inequality holds because the middle term is decreasing in n (since $1-(1-\frac{1}{n})^{n-2}$ is increasing in n), and the value of the middle term at $n = 50$ is less than 7.9. Recall that the probability that G has no perfect matching $\mathbb{P}\left[\bar{\mathfrak{m}}\right]$ is at most n^{-3}. When G has no perfect matching the rank-optimal assignment has average rank at most n. Thus, the expected average rank of the rank-optimal assignment is at most $7.9 + n^{-2}$. (We remark that this bound can be improved for sufficiently large n, as in the latter inequality, the middle term approaches $3 + 3\frac{e}{e-1}$ as n approaches infinity. Thus, for sufficiently large n, the bound R on the expected average rank improves to any constant larger than $3 + 3\frac{e}{e-1} \approx 7.746$.)

To prove the theorem, it remains to prove the next claim.

Claim 4.10. $\mathbb{E}\left[w\right] \leq 3n(1+\frac{1}{1-(1-\frac{1}{n})^{n-2}})$.

Proof. Equivalently, we will show that for an object $o \in O$, $\mathbb{E}\left[w_o\right] \leq 3(1+\frac{1}{1-(1-\frac{1}{n})^{n-2}})$. Let $\mathcal{M}$ be an $n \times n$ matrix such that, for every $a \in A$, $\mathcal{M}_{a,1}, \ldots, \mathcal{M}_{a,n}$ is the preference list of agent a where objects are ordered from the most to the least preferred. We will use the Principle of Deferred Decisions for the proof. The idea is that, rather than drawing the preference profile π (associated with the matrix $\mathcal{M}$) at once, we determine the elements of the matrix $\mathcal{M}$ one by one through a stochastic process, namely $\mathcal{P}$.

The process $\mathcal{P}$ starts from an empty matrix $\mathcal{M}$ (with the values of its entries undetermined) and determines the values of the entries of $\mathcal{M}$ one by one. At *round* i of the process, the values of the entries in column $i+3$ (i.e., $\mathcal{M}_{1,i+3}, \ldots, \mathcal{M}_{n,i+3}$) will be determined one by one. The process stops at the moment that object o appears $d = 3$ times in the matrix, or at the end of round $n-3$, whichever occurs first. Let s denote the round at which the process stops. We will complete the proof in two steps by (i) showing that $s + 3 \geq w_o$ and (ii) providing an upper bound on $\mathbb{E}\left[s\right]$.

Step (i): $s + 3 \geq w_o$. Every edge in $\widehat{\widetilde{E}}$ that is adjacent to o has weight of at most 3. To see why, recall that $\widehat{\widetilde{E}}$ contains an edge (a, o) for every agent a that ranks o third or better on her preference list.

Also, every edge in $\widehat{E}$ that is adjacent to o has weight of at most $s+3$. To see why, we recall the definition of $\widehat{E}$. We defined r_o to be the smallest integer greater than $d = 3$ such that there exists at least 3 agents who rank o (strictly) worse than 3 and no worse than r_o. If there is no such integer r_o, then we defined $r_o = n$. $\overset{\curvearrowleft}{E}$ contains an edge (o, a) for every agent a that ranks o better than 3 and no better than r_o. Thus, process $\mathcal{P}$ stops in round $r_o - 3$. That is, $r_o = s + 3$. By definition, the weight of every edge in $\overset{\curvearrowleft}{E}$ that is adjacent to o is at most r_o. Therefore, the weight of every edge in $\overset{\curvearrowleft}{E}$ that is adjacent to o is at most $s + 3$. This bound, together with the fact that the weights of the edges in $\overset{\curvearrowright}{E}$ are at most 3, concludes step (i).

Step (ii): An upper bound on $\mathbb{E}[s]$. Consider a round $t \geq 1$ in process $\mathcal{P}$ and an agent a. If agent a has not listed object o on her preference list in the previous rounds (i.e., if $o \notin \{\mathcal{M}_{a,4}, \ldots, \mathcal{M}_{a,t+2}\}$), then $\mathcal{P}$ sets $\mathcal{M}_{a,t+3} = o$ with probability $\frac{1}{n-t+1}$, which is at least $\frac{1}{n}$. Conditional on process $\mathcal{P}$ not having stopped, there are at most 2 agents who have listed object o on their list in previous rounds. Thus, the probability that during round t no agent lists object o on her list is at most $(1 - \frac{1}{n})^{n-2}$. Process $\mathcal{P}$ stops at a round s when at least 3 agents list object o, or by the end of round $s = n - 3$ if that event never occurs. Since the chance that no agent lists o during a round is at most $(1 - \frac{1}{n})^{n-2}$, then s is stochastically dominated by the sum of three iid geometric random variables with failure probabilities $(1 - \frac{1}{n})^{n-2}$. Thus, $\mathbb{E}[s] \leq \frac{3}{1-(1-\frac{1}{n})^{n-2}}$.

Finally, from steps (i) and (ii) it follows that

$$\mathbb{E}[w_o] \leq \mathbb{E}[s+3] \leq 3 + \frac{3}{1 - (1 - \frac{1}{n})^{n-2}},$$

which concludes the proof. $\square$

Remark 4.11. *The above analysis can also provide an asymptotic upper bound—in the sense of first-order stochastic dominance—on the distribution of the rank of the object assigned to an agent, as described next. Let μ denote a perfect matching of G chosen uniformly at random; if G has no perfect matching, which can occur with probability at most n^{-3}, then draw an assignment of agents to objects uniformly at random and denote it by μ. Define the random variable $\mathfrak{p}_a$ as the rank of the object assigned to agent a in μ. Also, define the random variable $\mathfrak{q}_o$ as the rank that agent $\mu(o)$ assigns to o on her list. Observe that $\mathbb{P}[\mathfrak{p}_a = i] = \mathbb{P}[\mathfrak{q}_o = i]$ for every $a \in A$, $o \in O$, and rank $i \in [n]$ (This holds because the expected number of agents who are assigned to their i-th choice in μ is equal to $\sum_{a' \in A} \mathbb{P}[\mathfrak{p}_{a'} = i] = n\mathbb{P}[\mathfrak{p}_a = i]$, and also equal to $\sum_{o' \in O} \mathbb{P}[\mathfrak{q}_{o'} = i] = n\mathbb{P}[\mathfrak{q}_o = i]$). Therefore, showing that $\mathfrak{q}_o$ is first-order stochastically dominated by some distribution D also implies that $\mathfrak{p}_a$ is first-order stochasti-*

cally dominated by D. When G has a perfect matching, $\mathfrak{q}_o \le w_o$. Step (ii) in the proof of Claim 4.10 proves that w_o is stochastically dominated by a random variable $Z = 3 + \sum_{j=1}^{3} Z_j$ where Z_j's are iid geometric random variables with failure probabilities $(1 - \frac{1}{n})^{n-2}$. Hence, $\mathbb{P}[\mathfrak{q}_o > i] \le n^{-3} + \mathbb{P}[Z > i]$; which also means that $\mathbb{P}[\mathfrak{p}_a > i] \le n^{-3} + \mathbb{P}[Z > i]$. This is the promised asymptotic upper bound.

5 Mechanisms for Eliciting Private Preferences

In this section we take a mechanism design perspective by considering the setting where every agent's preference list is her private information. After noting that there exists no *Dominant-Strategy Incentive-Compatible* (DSIC) mechanism that implements a rank-optimal assignment, we show that such mechanisms exist if the notion of incentive compatibility is relaxed to *Bayesian Incentive Compatiblity* (BIC).

Consider a market, as defined in our main setup in Section 3, with m objects each with capacity c and $n = mc$ agents. Given a preference profile π in this market, we use π_i to denote the preference list of agent i and the vector $\pi_{-i} = (\pi_j)_{j \in A, j \neq i}$ to denote the preference lists of the rest of the agents. Let Π_{-i} denote the set of all such possible π_{-i}. Given a preference list σ, we sometimes use $[\sigma; \pi_{-i}]$ to denote a preference profile π where $\pi_i = \sigma$ and π_{-i} determines the preference lists of the agents in $A \backslash \{i\}$. Finally, let M denote the set of all assignments in the market.

5.1 DSIC Mechanisms

A mechanism is a function $\mathcal{M} : \Pi \to M$. It elicits a preference list π_i from every agent i, constructs the preference profile $\pi = (\pi_1, \ldots, \pi_n)$, and then assigns objects to agents according to $\mathcal{M}(\pi)$. For every agent i, let $\mathcal{M}_i(\pi)$ denote the object assigned to i in the assignment $\mathcal{M}(\pi)$. A mechanism is DSIC if for every agent i, every preference order σ over O, and every π_{-i}, agent i weakly prefers the object $\mathcal{M}_i(\pi)$ to the object $\mathcal{M}_i([\sigma; \pi_{-i}])$.

A mechanism $\mathcal{M}$ is *rank-optimal* if, for every preference profile π, $\mathcal{M}(\pi)$ is a rank-optimal assignment (with respect to the objects' ranks in π). From a result of Featherstone (2020), it follows that no DSIC rank-optimal mechanism exists in a market with 4 agents and 4 objects each with capacity 1.[26] The proof is by constructing a specific market. Our next proposition shows that rank-optimal DSIC mechanisms exist in no parametrization of our

[26]His result is in fact stronger, in that he shows that no DSIC mechanism can be *rank efficient* (in the sense of first-order stochastic dominance) in the market that he constructs.

setup (i.e., choices of c, m) when $m \geq 4$. The proof is by constructing a (different) market for every such c, m.

Proposition 5.1. *No rank-optimal DSIC mechanism exists when $m \geq 4$.*

A well-known DSIC mechanism that is applicable to our setup is the Random Serial Dictatorship (RSD) mechanism (Abdulkadiroglu and Sönmez, 1998). RSD orders agents uniformly at random and then allows each agent to choose her most preferred object from the set of remaining objects, one agent at a time in that order. We call the assignment generated by RSD the *RSD assignment*. Remarkably, there is a quite large gap between the average ranks of the RSD and rank-optimal assignments. When $c = 1$, Knuth (1996) shows that the expected average rank of the RSD assignment almost equals $\ln n$, and thus approaches infinity as n does, in contrast to the average rank of the rank-optimal assignment.[27] We next show that the gap between the average ranks of these two assignments persists even when the object capacities are greater than one—in particular, when $c = o(\log n)$.[28]

Proposition 5.2. *The expected average rank of the RSD assignment is at least $\frac{\ln m - 1}{c}$.*

It remains an open question whether there exist DSIC mechanisms with a constant average rank (i.e., an average rank independent of the market size). We remark that the answer is negative if the mechanism is required to be *Pareto efficient*[29] and *nonbossy*[30] as well. This is because any such mechanism is equivalent to RSD, as shown by Bade (2020).

We next relax the notion of DSIC to BIC and ask whether rank-optimal BIC mechanisms exist in our main setup.

5.2 BIC Mechanisms in Uniform Markets

To formally define BIC mechanisms, we allow a mechanism to be a function $\mathcal{M} : \Pi \to \Delta(M)$; that is, $\mathcal{M}(\pi)$ is a probability distribution over assignments for every $\pi \in \Pi$.

For an assignment μ and an agent i, define $\mathbb{1}_i^\mu \in \{0, 1\}^m$ to be a vector v such that $v_o = 1$ if and only if $o = \mu(i)$. Let $\mathcal{M}_i(\pi) = \mathbb{E}_{\mu \sim \mathcal{M}(\pi)} [\mathbb{1}_i^\mu]$. The *interim allocation rule* η of

[27] We remark that if there is a linear excess of objects, i.e., $m = (1 + \epsilon)n$ for a constant $\epsilon > 0$, then the assignment generated by RSD has a constant average rank (e.g., see (Ashlagi and Nikzad, 2020), Lemma D.15).

[28] We recall that for two functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$, the notation $f = o(g)$ means $\lim_{n\to\infty} f(n)/g(n) = 0$.

[29] An assignment is Pareto efficient if no other assignment exists in which all agents are weakly better off with at least one agent being strictly better off.

[30] Nonbossyness essentially means that an agent can only change someone else's match if she also changes her own match. See Bade (2020) for the formal definitions and result.

a mechanism $\mathcal{M}$ is a sequence of functions $\eta_1, \ldots, \eta_n$ where, for every preference list σ and every agent i, $\eta_i(\sigma) = \mathbb{E}_{\pi_{-i} \sim U(\Pi_{-i})} [\mathcal{M}_i([\sigma; \pi_{-i}])]$.

A mechanism $\mathcal{M}$ with an interim allocation rule η is BIC if, for every agent $i \in A$ with preference order π_i, and every strict preference order σ over the set of objects, the following holds. Let $v = \eta_i(\pi_i)$ and $v' = \eta_i(\sigma)$; also, let $\preceq_i = \pi_i$;[31] then, for every $o \in O$

$$\sum_{q \in O : o \preceq_i q} v_q \geq \sum_{q \in O : o \preceq_i q} v'_q. \tag{5.1}$$

Let $\mathcal{M}^*$ be the mechanism in which, for every $\pi \in \Pi$, $\mathcal{M}^*(\pi)$ is the uniform distribution over all rank-optimal assignments. (Thus, if the rank-optimal assignment is unique for a given preference profile π, then $\mathcal{M}^*(\pi)$ is simply the unique rank-optimal assignment.)

Proposition 5.3. *The mechanism $\mathcal{M}^*$ is BIC.*

This shows the existence of rank-optimal BIC assignments in uniform markets (i.e., where the agents' preference orders are drawn iid and uniformly at random). The existence of rank-optimal BIC mechanisms in markets where the agents' preference orders are correlated or have a nonuniform distribution remains an interesting open question.

6 Conclusion

We show that in markets where agents rank objects uniformly at random, the average rank of the rank-optimal assignment (i.e., the first-best average rank) is bounded from above by a constant independent of the number of agents, and thus does not grow infinitely large as the number of agents does. At the same time, the average rank under the RSD mechanism can grow infinitely large when the object capacities grow at a sufficiently slow rate. Thus, the average rank can take a heavy toll under the RSD mechanism compared to the first-best. Furthermore, while no DSIC mechanism can implement the rank-optimal assignment when the agents' rankings are their private information, we show that there exists a BIC mechanism that does so. These findings highlight the possibility of using assignment methods other than RSD in scenarios where the average rank is a relevant objective.

[31] Hence, $o_1 \preceq_i o_2$ means that object o_1 is ranked weakly lower than o_2 on the preference list π_i.

References

Atila Abdulkadirouglu, Parag A Pathak, and Alvin E Roth. 2005. The New York City High School Match. *American Economic Review* 95, 2 (2005), 364–367.

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

Itai Ashlagi, Yash Kanoria, and Jacob D. Leshno. 2017. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy* 125, 1 (2017), 69–98.

Itai Ashlagi and Afshin Nikzad. 2020. What matters in school choice tie-breaking? How competition guides design. *Journal of Economic Theory* 190 (2020), 105120. https://doi.org/10.1016/j.jet.2020.105120

Itai Ashlagi, Afshin Nikzad, and Assaf Romm. 2019. Assigning more students to their top choices: A comparison of tie-breaking rules. *Games and Economic Behavior* 115 (2019), 167 – 187. https://doi.org/10.1016/j.geb.2019.02.015

Sophie Bade. 2020. Random Serial Dictatorship: The One and Only. *Mathematics of Operations Research* 45, 1 (2020), 353–368. https://doi.org/10.1287/moor.2019.0987 arXiv:https://doi.org/10.1287/moor.2019.0987

F. T. Bruss. 2000. Sum the odds to one and stop. *Annals of Probability* 28 (06 2000). https://doi.org/10.1214/aop/1019160340

Eric Budish and Estelle Cantillon. 2012. The multi-unit assignment problem: Theory and evidence from course allocation at Harvard. *American Economic Review* 102, 5 (2012), 2237–71.

Yeon-Koo Che and Olivier Tercieux. 2018. Payoff equivalence of efficient mechanisms in large matching markets. *Theoretical Economics* 13, 1 (2018), 239–271. https://doi.org/10.3982/TE2793 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE2793

Young H. Chun and Robert T. Sumichrast. 2006. A rank-based approach to the sequential selection and assignment problem. *European Journal of Operational Research* 174, 2 (2006), 1338 – 1344. https://doi.org/10.1016/j.ejor.2005.03.040

Shagnik Das. 2016. A brief note on estimates of binomial coefficients. http://page.mi.fu-berlin.de/shagnik/notes/binomials.pdf

Monique De Haan, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw. 2015. The performance of school assignment mechanisms in practice. *CEPR Discussion Paper No. DP10656* (2015).

Clayton R Featherstone. 2020. Rank Efficiency: Modeling a Common Policymaker Objective. *Unpublished paper, The Wharton School, University of Pennsylvania.[252]* (2020).

C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. 1971. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics* 22, 2 (01 Jun 1971), 89–103. https://doi.org/10.1007/BF01651330

Nikolaos Fountoulakis and Konstantinos Panagiotou. 2012. Sharp load thresholds for cuckoo hashing. *Random Structures & Algorithms* 41, 3 (2012), 306–333.

Alan Frieze and Michal Karonski. 2012. Introduction to Random Graphs. (2012). https://www.math.cmu.edu/~af1p/BOOK.pdf

Alan Frieze and Páll Melsted. 2009. Maximum matchings in random bipartite graphs and the space utilization of Cuckoo Hash tables. *Random Structures & Algorithms* 41, 3 (2009), 334–364. https://doi.org/10.1002/rsa.20427 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20427

Donald E. Knuth. 1996. An Exact Analysis of Stable Allocation. *Journal of Algorithms* 20, 2 (1996), 431 – 442. https://doi.org/10.1006/jagm.1996.0020

SangMok Lee and Leeat Yariv. 2018. On the Efficiency of Stable Matchings in Large Markets. *Working paper* (2018). https://arxiv.org/abs/1512.07276

Svante Linusson and Johan Wästlund. 2004. A proof of Parisi's conjecture on the random assignment problem. *Probability Theory and Related Fields* 128, 3 (01 Mar 2004), 419–440. https://doi.org/10.1007/s00440-003-0308-9

Chandra Nair, Balaji Prabhakar, and Mayank Sharma. 2005. Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures. *Random Structures & Algorithms* 27, 4 (2005), 413–444. https://doi.org/10.1002/rsa.20084 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20084

NRMP. 2020. Charting Outcomes in the Match: Senior Students of U.S. MD Medical Schools. https://therideshareguy.com/what-its-like-to-be-part-of-lyfts-new-prime-time-experiment/.

Boris Pittel. 1989. The average number of stable matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 530–549.

Boris Pittel. 1992. On likely solutions of a stable marriage problem. *The Annals of Applied Probability* (1992), 358–401.

Lloyd Shapley and Herbert Scarf. 1974. On cores and indivisibility. *Journal of mathematical economics* 1, 1 (1974), 23–37.

David W. Walkup. 1980. Matchings in random regular bipartite digraphs. *Discrete Mathematics* 31, 1 (1980), 59 – 64. https://doi.org/10.1016/0012-365X(80)90172-7

Douglas B. West. 2000. *Introduction to Graph Theory* (2 ed.). Prentice Hall.

A Proofs from Section 4

Consider a market with set of agents A and objects O where each agent has a complete strict preference list over objects. A *preference profile* is a function $\pi : A \to O^{|A|}$ that determines the preference list of each agent. We use a more compact notation of π_a instead of $\pi(a)$ to denote the preference list of an agent a.

Proof of Fact 4.3. We will show that the lattice defined by the relation $\preceq_l$ is the Cartesian product of two lattices, namely $\mathcal{L}^1, \mathcal{L}^2$. The claim then would be proved since the Cartesian product of two lattices is a lattice itself.

The set of the elements of the lattice $\mathcal{L}^1$ is denoted by V_l^1. An element $w \in V_l^1$ is a sequence $(w^1, \dots, w^k)$ where w^i is a subset of $[n]$ with size l. Similarly, the set of the elements of $\mathcal{L}^2$ is V_l^2. An element $w \in V_l^2$ is a sequence $(w^{k+1}, \dots, w^n)$ where w^i is a subset of $[n]$ with size l.

Recall that $h(i) = i$ if $i > d$ and let $h(i) = n + i$ otherwise. Also, recall that for integers α, β, we defined $\alpha \preceq \beta$ if $h(\alpha) \leq h(\beta)$. Similarly, for two vectors of the same length u, v, $u \preceq v$ if $u_j \preceq v_j$ for every coordinate j.

For $w, z \in V_l^1$, we define $w \preceq_l^1 z$ when $\overrightarrow{z^a} \preceq \overrightarrow{w^a}$ for all $a \in X$. We observe that $\preceq_l^1$ is a lattice, because it has a well-defined meet operator corresponding to coordinate-wise maximum with respect to $\preceq$, and a well-defined join operator corresponding to coordinate-wise minimum with respect to $\preceq$.

For $w, z \in V_l^2$, define $w \preceq_l^2 z$ if $\overrightarrow{w^a} \preceq \overrightarrow{z^a}$ for all $a \in A \backslash X$. The relation $\preceq_l^2$ is a lattice, because it has a well-defined meet operator corresponding to coordinate-wise minimum with respect to $\preceq$, and a well-defined join operator corresponding to coordinate-wise maximum with respect to $\preceq$.

Consider $x, y \in V^l$ with $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$. By the definition of $\preceq_l$, we have $x \preceq_l y$ if

$$(x^1, \ldots, x^k) \preceq_l^1 (y^1, \ldots, y^k), \tag{A.1}$$

$$(x^{k+1}, \ldots, x^n) \preceq_l^2 (y^{k+1}, \ldots, y^n). \tag{A.2}$$

That is, the relation $\preceq_l$ is the Cartesian product of the lattices $\preceq_l^1$ and $\preceq_l^2$. Thus, $\preceq_l$ itself is a lattice. $\square$

Proof of Fact 4.4. Recall that R denotes the event $(X \times Y) \cap \widehat{E} = \varnothing$ where $\widehat{E}$ contains an edge (a, o) for every agent a and every object o that is ranked d or better on a's preference list. Consider an arbitrary preference profile $\pi \in \Pi$ and recall that $\mathcal{L}_\pi = (\mathcal{L}_\pi^1, \ldots, \mathcal{L}_\pi^n)$ where, for every $a \in A$, $\mathcal{L}_\pi^a$ is the set of ranks of objects $1, \ldots, l$ on agent a's preference list. We observe that $\pi \in R$ if and only if $\mathcal{L}_\pi^a \cap \{1, \ldots, d\} = \varnothing$ for every $a \in X$. Therefore, $f_R(\mathcal{L}_\pi) = 1$ if the latter condition holds; otherwise, $f_R(\mathcal{L}_\pi) = 0$. Thus, $f_R(\mathcal{L}_\pi) \in \{0, 1\}$ for every $\pi \in \Pi$. This proves the first claim since, for every $x \in V_l$, there exists $\pi \in \Pi$ such that $\mathcal{L}_\pi = x$.

To prove the second claim, consider preference profiles π, π' such that $\mathcal{L}_\pi \preceq_l \mathcal{L}_{\pi'}$. For every agent $a \in X$, $\mathcal{L}_\pi^a \cap \{1, \ldots, d\} = \varnothing$ implies that $\mathcal{L}_{\pi'}^a \cap \{1, \ldots, d\} = \varnothing$, because $\mathcal{L}_\pi \preceq_l \mathcal{L}_{\pi'}$. Thus, $f_R(\mathcal{L}_\pi) = 1$ implies that $f_R(\mathcal{L}_{\pi'}) = 1$, which proves that $f_R(x) \leq f_R(y)$.

It remains to prove that $f_L(x) \geq f_L(y)$. We need a few definitions first. An *extension* of an element $x = (x^1, \ldots, x^n)$ of the lattice, namely χ, is defined as follows. Recall that, for each $a \in A$, x^a is subset of $[n]$ of size l. Then, we define $\chi = (\chi^1, \ldots, \chi^n)$, where χ^a is an l-dimensional vector that contains the elements of x^a in some order. (Thus, every element x of the lattice has $(l!)^n$ distinct extensions.) The interpretation of χ is as follows. Suppose $\chi^a = (\chi^{a,1}, \ldots, \chi^{a,l})$. Then, $\chi^{a,j}$ is the rank of object j on the agent a's preference list.

We say a preference profile π is *aligned* with the extension χ if, for every agent $a \in A$ and object $j \in Y$, $\chi^{a,j}$ equals the rank of object j on agent a's preference list in π. (Thus, are $((n-l)!)^n$ distinct preference profiles that are aligned with χ, as the objects in $O \backslash Y$ can be ranked in $(n-l)!$ different ways on each agent's preference list.)

An extension χ is *consistent* with event L if there exists a preference profile π aligned with χ such that $\pi \in L$. (We observe that if there exists one such preference profile π, then for any preference profile π' aligned with χ, $\pi' \in L$. This holds because only the ranks of the objects in $Y = \{1, \ldots, l\}$ on the agents' lists determine whether a preference profile belongs to L, and these ranks are determined by χ.)

Let Ξ_x denote the number of extensions of x consistent with L. Therefore,

$$f_L(x) = \frac{\Xi_x \left((n-l)!\right)^n}{(l!)^n((n-l)!)^n} = \frac{\Xi_x}{(l!)^n}.$$

To see why this equality holds, observe that in the middle term (i) the numerator is the total number of preference profiles belonging to $\Pi_x \cap L$, and (ii) the denominator is the total number of preference profiles in Π_x.

By the above equation, to prove $f_L(x) \geq f_L(y)$, it suffices to show that $\Xi_x \geq \Xi_y$. We will prove this by constructing a bijection B from the set of extensions of y to the set of extensions of x such that for every extension ψ of y, $B(\psi)$ is consistent with L if ψ is consistent with L.

Recall that $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$ where x^a and y^a are subsets of $[n]$ with size l. For every a, let $\overrightarrow{x^a} = (x^{a,1}, \ldots, x^{a,l})$ and $\overrightarrow{y^a} = (y^{a,1}, \ldots, y^{a,l})$. Let C^a be a bijection from y^a to x^a such that $C^a(y^{a,j}) = x^{a,j}$ for every $j \in [l]$. For an extension $\psi = (\psi^1, \ldots, \psi^n)$ of y with $\psi^a = (\psi^{a,1}, \ldots, \psi^{a,l})$ for every $a \in A$, define $B(\psi)$ to be the extension $\chi = (\chi^1, \ldots, \chi^n)$ with

$$\chi^a = (C^a(\psi^{a,1}), \ldots, C^a(\psi^{a,l})).$$

Since $x \preceq y$, then $y^{a,j} \preceq C^a(y^{a,j})$ for $a \in X$, and $C^a(y^{a,j}) \preceq y^{a,j}$ for $a \notin X$. Therefore, if ψ is consistent with L, $B(\psi)$ is also consistent with L. This proves the promised claim. $\square$

Proof of Fact 4.6. The proof for f follows from Fact 4.4, as described next. In Fact 4.4 we showed that $f_L(x) \geq f_L(y)$ for $x \preceq_l y$, where

$$f_L(x) = \mathbb{P}_{\pi \sim U(\Pi_x)} \left[\pi \in L_1 \cap \ldots \cap L_l\right].$$

The function f here is the same as the function f_L, with the difference that f is defined over the lattice $\mathcal{L}[l']$ whereas f_L is defined over $\mathcal{L}[l]$. Thus, the proof for $f(x) \geq f(y)$ if $x \preceq_{l'} y$ is the same as the proof for $f_L(x) \geq f_L(y)$ if $x \preceq_l y$ but for l replaced with l'.

To prove the claim for g, we recall the lattice $\mathcal{K} = \mathcal{L}[l']$ defined over the set $V_{l'}$. Also, recall that we say a preference profile π is associated with an element of the lattice, denoted by $\mathcal{K}_\pi = (\mathcal{K}_\pi^1, \ldots, \mathcal{K}_\pi^n)$ where $\mathcal{K}_\pi^a$ is the set of ranks of the objects $1, \ldots, l'$ on the agent a's preference list. We let $\Pi_x = \{\pi \in \Pi : x = \mathcal{K}_\pi\}$.

Recall that $g(z) = \mathbb{P}_{\pi \sim U(\Pi_z)}[\pi \in L_{l'+1}]$ for every $z \in V_{l'}$. Thus, to prove that $g(x) \le g(y)$, it suffices to prove the following: there exists a bijection B from Π_x to Π_y such that if $\pi \in L_{l'+1}$ for a preference profile $\pi \in \Pi_x$, then $B(\pi) \in L_{l'+1}$. We construct B as follows.

Let $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$ be the lattice elements. Consider $\pi \in \Pi_x$. Let $r_a^\pi(o)$ denote the rank of object o on a's preference list in the preference profile π. We define $B(\pi)$ to be the unique preference profile π' in which, for every agent $a \in A$ it holds that[32]

i. for all objects $o_1, o_2 \in [l']$, $r_a^\pi(o_1) \preceq r_a^\pi(o_2)$ if and only if $r_a^{\pi'}(o_1) \preceq r_a^{\pi'}(o_2)$ and

ii. for all objects $o_1, o_2 \notin [l']$, $r_a^\pi(o_1) \preceq r_a^\pi(o_2)$ if and only if $r_a^{\pi'}(o_1) \preceq r_a^{\pi'}(o_2)$.

We note that the set of ranks $\{r_a^\pi(o) : o \in [l']\}$ and $\{r_a^{\pi'}(o) : o \in [l']\}$ are determined by x^a and y^a, respectively. Thus, the above two conditions uniquely determine the preference list of every agent a in $\pi' = B(\pi)$. Figure 5 demonstrates the construction of π' by an example.

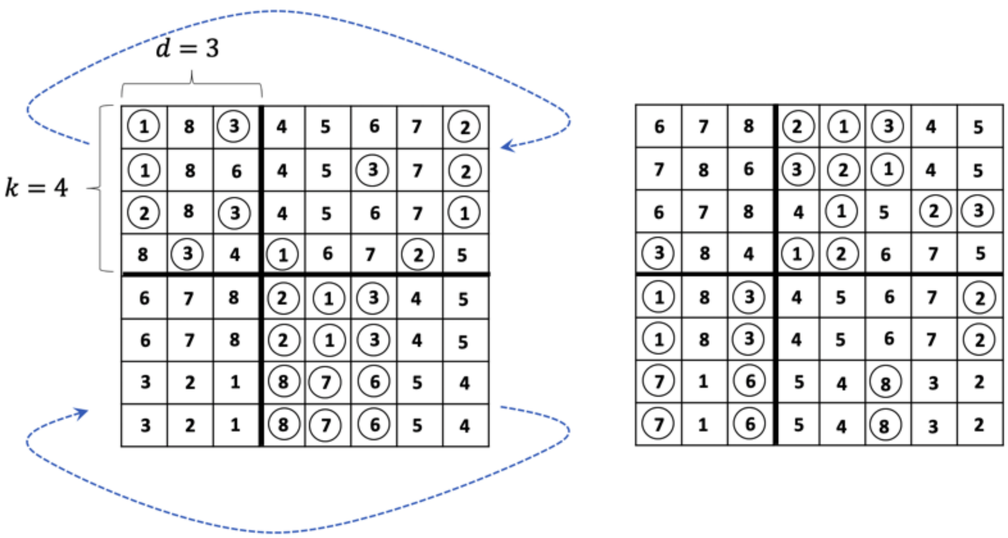

Figure 5: $n = 8$, $k = 4$, and $l' = 3$. The left and right panels correspond to $\pi \in \Pi_x$ and $\pi' \in \Pi_y$, respectively. The rows correspond to the preference lists of agents 1 to n, from top to bottom, respectively. On each list the objects are ordered from left to right in the order of the most favorite to the least favorite. We have $x^1 = \{1, 3, 8\}$ (these positions are marked with a circle in the first row of π), $x^2 = \{1, 6, 8\}, x^3 = \{1, 3, 8\}, x^4 = \{2, 4, 7\}, x^5 = x^6 = x^7 = x^8 = \{4, 5, 6\}$. Also, $y^1 = y^2 = \{4, 5, 6\}, y^3 = \{5, 7, 8\}, y^4 = \{1, 4, 5\}, y^5 = y^6 = \{1, 3, 8\}$, and $y^7 = y^8 = \{1, 3, 6\}$.

By the construction of π', $r_{\pi'}^a(l'+1) \preceq r_\pi^a(l'+1)$ for $a \in [k]$ and $r_\pi^a(l'+1) \preceq r_{\pi'}^a(l'+1)$ for $a \notin [k]$ hold since $x \preceq_{l'} y$. This implies that if $\pi \in L_{l'+1}$, then $\pi' \in L_{l'+1}$. That is, if

[32]For the following definition, we recall that for integers α, β, we defined $\alpha \preceq \beta$ if $h(\alpha) \le h(\beta)$, where $h(i) = i$ if $i > d$ and $h(i) = n + i$ otherwise.

$\pi \in L_{l'+1}$, then $B(\pi) \in L_{l'+1}$, which is the promised claim. $\square$

Proof of Fact 4.7. Recall that x is a sequence $(x^1, \ldots, x^n)$ where, for every $a \in A$, x^a is the set of ranks of objects $1, \ldots, l$ on agent a's preference list. For a subset $Q \subseteq O$, a *partial preference profile* $r^Q : A \times Q \to [n]$ is a function such that $r^Q(a, o) \neq r^Q(a, o')$ for every distinct $o, o' \in Q$ and $a \in A$. Given a preference profile π and a subset $Q \subseteq O$, we let π^Q denote a partial preference profile such that $\pi^Q(a, o)$ denotes the rank of an object $o \in Q$ on an agent a's preference list.

Let $Y' = \{1, \ldots, l'\}$. Since $\pi \in \Pi_x$, then, an object in Y' has a rank belonging to the set x^a, and object $l'+1$ has a rank belonging to the set $[n] \backslash x^a$ on the preference list of an agent a. Thus, when $\pi \sim U(\Pi_x)$, conditioning on $\pi^{\{l'+1\}}$ does not change the distribution of $\pi^{Y'}$ (compared to the unconditional distribution of $\pi^{Y'}$). Since $\pi^{Y'}$ and $\pi^{\{l'+1\}}$ are respectively sufficient statistics for determining whether the events $\pi \in L_1 \cap \ldots \cap L_{l'}$ and $\pi \in L_{l'+1}$ hold, then these events are independent when $\pi \sim U(\Pi_x)$. $\square$

A.1 Proof for the Case of Non-Unit Capacities

The proof for non-unit capacities ($c > 1$) is a corollary of the proof for unit capacities. We first need a few definitions. The *rank distribution* of an assignment $\mu : A \to O$ is a vector $\mathcal{R}_\mu \in \mathbb{R}_+^{|O|}$ where its i-th entry $\mathcal{R}_\mu(i)$ denotes the fraction of agents who are assigned by μ to their i-th favorite object (i.e., the object that they rank i).

Consider a uniform market M with its preference profile π drawn uniformly at random from the set of all preference profiles possible in M. Let μ_π denote the rank-optimal assignment in M. The *expected* rank distribution of the rank-optimal assignment in M is the vector $\mathbb{E}_\pi [\mathcal{R}_{\mu_\pi}]$.

We say a vector $\mathcal{R}' \in \mathbb{R}_+^{n_1}$ with $\sum_{i=1}^{n_1} \mathcal{R}'(i) = 1$ *rank-wise* dominates another vector $\mathcal{R} \in \mathbb{R}_+^{n_2}$ with $\sum_{i=1}^{n_2} \mathcal{R}(i) = 1$ if for all positive integers $r \leq \min\{n_1, n_2\}$,

$$\sum_{i=1}^{r} \mathcal{R}'(i) \geq \sum_{i=1}^{r} \mathcal{R}(i).$$

We will use the next lemma to prove Theorem 3.1 for the case of $c > 1$.

Lemma A.1. *Let M, M' be two uniform markets both with n agents. Suppose that M has n objects each with capacity 1 and M' has n/c objects each with capacity $c \geq 1$, respectively. Let $\mathcal{R}, \mathcal{R}'$ denote the expected rank distributions of the rank-optimal assignments in M, M', respectively. Then, $\mathcal{R}'$ rank-wise dominates $\mathcal{R}$.*

Proof. Let r_π denote the average rank of the rank-optimal assignment in a market with preference profile π. Let Π, Π' denote the set of possible preference profiles in M, M'. The proof works by defining a function $f : \Pi \to \Pi'$ that maps every $\frac{|\Pi|}{|\Pi'|}$ elements of Π to precisely one element of Π. Moreover, this function is defined such that, for any $\pi \in \Pi$, the rank distribution of the rank-optimal assignment for π is rank-wise dominated by the rank distribution of the rank-optimal assignment for $f(\pi)$. To prove the lemma, it suffices to show that such a function exists as follows.

Let n be the number of agents in both markets. In the market M, relabel the objects as $O = \{o_t^j : j \in [c], t \in [n/c]\}$. We say an object o is of type t if $o = o_t^j$ for some j, t. With slight abuse of notation, we use $t(o)$ to denote the type of an object o. Given a preference list σ over O, define $\overline{\sigma}$ to be a list in which $\overline{\sigma}(i) = t(\sigma(i))$. (That is, each object is replaced with its type.)

Let $g(\sigma)$ be a preference order defined over $[n/c]$ as follows: in its i-th position, $g(\sigma)$ contains the i-th distinct number that appears in $\overline{\sigma}$, for $i \in [n/c]$. In other words, the function g removes the second, third, and higher appearances of a number in $\overline{\sigma}$ and outputs the resulting list.

For a preference profile $\pi \in \Pi$, define $f(\pi)$ to be the preference profile π' where $\pi'_a = g(\overline{\pi_a})$ for all $a \in A$. Observe that, by symmetry, $|f^{-1}(\pi')|$ does not depend on π', i.e., $|f^{-1}(\pi')| = \frac{|\Pi|}{|\Pi'|}$. To complete the proof, it remains to show that the rank distribution of the rank-optimal assignment for π is rank-wise dominated by the rank distribution of the rank-optimal assignment for π'. Let μ be the rank-optimal assignment for the preference profile π. We define the assignment μ' in the market M' for the preference profile π' as follows: for each agent a, let $\mu'(a) = t(\mu(a))$. Observe that μ' is a feasible assignment, and that $\mu'(a)$ does not have a worse rank on π'_a than $\mu(a)$ has on π_a. Therefore, the rank distribution of μ' rank-wise dominates the rank distribution of μ. This completes the proof. □

By the rank-wise dominance relation established by the above lemma, the expected average rank of the rank-optimal assignment is weakly higher in M than in M'. Thus, the upper bound on the expected average rank of the rank-optimal assignment for the case of $c = 1$ is a valid upper bound for the general case ($c \geq 1$) as well.

B Proofs from Section 5

Proof of Proposition 5.1. For a preference profile π and an assignment μ, let $R_\pi(\mu)$ denote the sum of the agents' ranks in the assignment μ, where the ranks are defined with respect to π.

We first construct the following market:

i. There are n agents, namely agents $1, \ldots, n$. Also, there are m objects, namely $o_1, \ldots, o_m$. All of the objects $o_1, \ldots, o_m$ have capacity c. The preference order of agent i is denoted by $\prec_i$.

ii. The preference orders of agents $1, \ldots, 4$ are defined by

$$o_1 \prec_1 o_2 \prec_1 o_3 \prec_1 o_4 \prec_1 o_5 \prec_1 \ldots \prec_1 o_m,$$
$$o_2 \prec_2 o_3 \prec_2 o_1 \prec_2 o_4 \prec_2 o_5 \prec_2 \ldots \prec_2 o_m,$$
$$o_3 \prec_3 o_2 \prec_3 o_4 \prec_3 o_1 \prec_3 o_5 \prec_3 \ldots \prec_3 o_m,$$
$$o_3 \prec_4 o_4 \prec_4 o_1 \prec_4 o_2 \prec_4 o_5 \prec_4 \ldots \prec_4 o_m.$$

iii. Agents $S_1 = \{5, \ldots, c+3\}$ rank object o_1 first and object o_4 last.[33]

iv. Agents in the set $S_2 = \{c+4, \ldots, 2c+2\}$ rank object o_2 first and object o_4 last.

v. Agents in the set $S_3 = \{2c+3, \ldots, 3c+1\}$ rank object o_3 first and object o_4 last.

vi. Agents in the set $S_4 = \{3c+2, \ldots, 4c\}$ rank object o_4 first.

vii. For every integer $i \in [5, m]$, every agent in the set $S_i = \{c(i-1)+1, \ldots, ci\}$ ranks object o_i first and ranks the rest of the objects in an arbitrary way in the remaining positions on her preference list such that objects o_1, o_2, o_3, o_4 are ranked as her least favorite objects, with o_4 being ranked last.

Denote the resulting preference profile by π. We first observe that there exists an assignment μ^* with $R_\pi(\mu^*) = n+1$: μ^* is defined by $\mu^*(j) = o_j$ for $j \in [4]$, and $\mu^*(a) = o_k$ for every agent $a \in S_k$ and $k \in [m]$. In fact, μ^* is the unique assignment satisfying $R_\pi(\mu^*) \leq n+1$. This holds because o_4 has capacity c, but only $c-1$ agents (i.e., those in S_4) rank o_4 first. On the other hand, the only agent not in S_4 who ranks o_4 second is agent 4. This uniquely determines μ^* as specified above.

We next suppose that agent 4 deviates by reporting

$$o_3 \prec_4 o_2 \prec_4 o_1 \prec_4 o_4 \prec_4 o_5 \prec_4 \ldots \prec_4 o_n.$$

We will show that, under this deviation, agent 4 will be assigned to object o_3, her top-ranked object.

Let the preference profile resulting from the deviation of agent 4 be denoted by $\bar{\pi}$. There

[33] The agents' preferences are otherwise arbitrary. For example, $o_1, o_2, o_3, o_5, \ldots, o_m, o_4$ could be such a preference order, listing objects from the most to the least preferred.

is no assignment μ' with $R_{\bar{\pi}}(\mu') < n+2$. This holds because precisely c agents should be assigned to o_4. Thus, at least one agent not in S_4 would be assigned to o_4. Note, however, that any such agent would rank o_4 third or worse in $\bar{\pi}$. Therefore, $R_{\bar{\pi}}(\mu') \geq n+2$ for every assignment μ'.

We next show that there exists a unique assignment μ with $R_{\bar{\pi}}(\mu) = n+2$. Recall that at least one agent not in S_4 would be assigned to o_4. The only agent not in S_4 who ranks o_4 third or better is agent 3, who ranks o_4 third. Therefore, $\mu(3) = o_4$ must hold if $R_{\bar{\pi}}(\mu) = n+2$. In addition, every agent other than 3 should be assigned to her top choice in such μ. This uniquely determines the assignment μ as follows: $\mu(1) = o_1$, $\mu(2) = o_2$, $\mu(3) = o_4$, $\mu(4) = o_3$, and $\mu(a) = o_k$ for every agent $a \in S_k$ and $k \in [m]$. We thus have shown that μ is the unique rank-optimal assignment in $\bar{\pi}$.

Finally, we note that $\mu^*(4) \prec_4 \mu(4)$. Therefore, agent 4 can strictly improve her assigned object by misreporting her preference list. This completes the proof. $\square$

Proof of Proposition 5.3. Consider an arbitrary agent $a \in A$. For every strict preference order σ over O and every $r \in [m]$, let $p_\sigma(r)$ denote the probability that agent a is assigned to the object ranked r-th on σ if she reports σ to $\mathcal{M}^*$. (We note this probability is computed by taking into account the randomization used by the mechanism as well as the randomization of π_{-a}.) Since the agents' preference lists are drawn iid and uniformly at random, then $p_\sigma(r) = p_{\sigma'}(r)$ for every preference list σ'. We therefore use $p(r)$ to denote the latter quantity, with slight abuse of notation.

We next show that $p(r)$ is decreasing in r. Suppose this is not the case. Then, there exists s such that $p(s) < p(s+1)$. Consider the mechanism $\mathcal{M}'$ that, after receiving the agents' preferences, (i) for every agent it swaps the objects at positions s and $s+1$ on her preference list and then (ii) runs $\mathcal{M}^*$ on the resulting preference profile. Let the random variables μ^* and μ' respectively denote the assignments generated by $\mathcal{M}^*$ and $\mathcal{M}'$. Since $p(s) < p(s+1)$, then the average rank of μ' is smaller than the average rank of μ, in expectation. This, however, contradicts the fact that $\mathcal{M}^*$ selects only rank-optimal assignments. Thus, $p(s)$ is decreasing in s. This implies that no agent can benefit from misreporting her preference list, in the sense of (5.1). $\square$

Proof of Proposition 5.2. Consider the last m agents who choose (i.e., the m agents with lowest priority numbers). Let them be indexed by $a_0, \ldots, a_{m-1}$, ordered with respect to their priority numbers with a_0 having the best priority number and a_{m-1} having the worst. Also, let R_i denote the expected rank of agent a_i and $R = \frac{1}{m} \cdot \sum_{i=0}^{m-1} R_i$.

To provide a lower bound on R_i, we define an auxiliary problem instance, which is running RSD on a market with m agents, namely $a'_0, \ldots, a'_{m-1}$ and m objects with unit capacities. Suppose the agents rank objects independently and uniformly at random and that agents choose objects in the same order as their indices: agent a'_0 chooses first. Let R'_i denote the average rank of agent a'_i, and $R' = \frac{1}{m} \cdot \sum_{i=0}^{m-1} R'_i$. Knuth (1996) shows that $R' \geq \ln m - 1$. The next claim states that $R_i \geq R'_i$, which implies that $R \geq R'$. That would complete the proof as it shows that the expected average rank in the original problem is at least $R' \cdot \frac{m}{n} \geq \frac{\ln m - 1}{c}$.

Claim B.1. *For any $i \in \{0, \ldots, m-1\}$, $R_i \geq R'_i$.*

Proof. When agent a'_i is choosing in the auxiliary problem, there are exactly i objects allocated by the agents before her, and therefore $m - i$ possible choices remain. In the original problem, when agent a_i is choosing, at most i objects have not run out of capacity; therefore, agent a_i has at most $m - i$ possible choices. This implies that the rank distribution for agent a'_i rank-wise dominates the rank distribution for agent a_i, which implies that $R_i \geq R'_i$. $\square$

$\square$

Online Appendix

i Markets with a Shortage or Excess of Objects

We show that Theorem 3.1 holds also when there is a shortage or excess of objects. In these cases, we consider the average rank in assignments that assign precisely $\min\{m, n\}$ agents to objects, where the average rank is defined as the sum of the ranks of the assigned agents normalized by the number of the assigned agents.

First we suppose there is a shortage of objects, i.e., $mc < n$. Consider an auxiliary uniform market in which $n - mc$ agents are removed from the original market. Thus, the auxiliary market has $n' = mc$ agents. By Theorem 3.1, the expected average rank of the rank-optimal assignment in the new market is at most R, for sufficiently large n'. On the other hand, for any realization of the preference profile in the original market, the average rank of the rank-optimal assignment in the original market is smaller (i.e., better) than in the auxiliary market, because some of the agents are excluded from the auxiliary market. Thus, the expected average rank in the original market is at most R.

We now consider the case where there is an excess of objects. First, we note that Lemma A.1 applies here as well, by the same proof. Thus, it suffices to prove the claim for $c = 1$. The proof considers two cases: either $m > 2n$ or not.

Case 1: $m > 2n$. In this case, we show that the assignment generated by the RSD mechanism has a constant average rank. Without loss of generality, suppose that agents $1, \dots, n$ choose objects one by one in that order. Since the market is uniform, agent a attains her top-ranked object with probability $\frac{m-a+1}{m} > 1/2$. Conditional on not attaining her top-ranked object, she will attain her second-ranked object with probability at least $\frac{m-a+1}{m-1} > 1/2$. Similarly, for $i \geq 2$, conditional on not attaining her i-th ranked object, she would attain her $(i+1)$-th ranked object with probability at least $\frac{m-a+1}{m-i+1} > 1/2$. Therefore, the rank of agent a is stochastically dominated by a geometric random variable with success probability $1/2$. Thus, her rank, in expectation, is at most 2. This proves the claim in Case 1.

Case 2: $n < m \leq 2n$. Let $O = \{o_1, \dots, o_m\}$ be the set of objects. In this case, we consider an auxiliary market that contains the same set of agents, A, as in the original market, but contains only a subset $O' = \{o_1, \dots, o_n\}$ of the objects.

Given that π denotes the preference profile of the original market, the preference profile π' in the auxiliary market is taken to be the same as π but with the objects in $O \backslash O'$ removed from the agents' lists (without changing the relative ordering of the rest of the objects). We note that π is a random variable drawn uniformly at random from the set of all preference profiles in the original market and that π' is uniquely determined by π.

Let $r^\star(\pi)$ and $r^\star(\pi')$ denote the average ranks of the rank-optimal assignments in the original and in the auxiliary market, respectively.

Lemma i.1. $\mathbb{E}_\pi [r^\star(\pi)] \leq 2\mathbb{E}_\pi [r^\star(\pi')]$.

Proof. Let $\mu_\pi, \mu_{\pi'}$ denote the rank-optimal assignments in the original market with preference profile π and the auxiliary market with preference profile π'. Also, let $r^\star_a(\pi)$ and $r^\star_a(\pi')$ respectively denote the ranks of a in μ_π and $\mu_{\pi'}$. To prove the claim, it suffices to show that $\mathbb{E}_\pi [r^\star_a(\pi)] \leq 2\mathbb{E}_\pi [r^\star_a(\pi')]$ holds for every agent $a \in A$.

To this end, let $\preceq^\pi_a$ denote the preference order of a in π. We first observe that

$$r^\star_a(\pi) = r^\star_a(\pi') + |\{o \in O \backslash O' : \mu_{\pi'}(a) \preceq^\pi_a o\}|.$$

Since the agents' preference orders are drawn independently and uniformly at random, then for every π' and $o \in O \backslash O'$,

$$\mathbb{P}_{\pi}\left[\mu_{\pi'}(a) \preceq_a^{\pi} o \mid \pi'\right] = \frac{r_a^{\star}(\pi')}{n+1}.$$

The two latter equations then imply that

$$\mathbb{E}_{\pi}\left[r_a^{\star}(\pi) \mid \pi'\right] = \sum_{o \in O \backslash O'} \mathbb{P}_{\pi}\left[\mu_{\pi'}(a) \preceq_a^{\pi} o \mid \pi'\right] = r_a^{\star}(\pi') + \frac{r_a^{\star}(\pi')}{n+1}|O \backslash O'| < 2r_a^{\star}(\pi'),$$

where the inequality holds because $|O \backslash O'| \leq n$. Applying the law of iterated expectations to the left- and the right-hand side proves the promised claim $\mathbb{E}_{\pi}\left[r_a^{\star}(\pi)\right] \leq 2\mathbb{E}_{\pi}\left[r_a^{\star}(\pi')\right]$. $\square$

ii Failure of Alternative Approaches

We next sketch two natural but unsuccessful proof approaches based on the Parisi conjecture and discuss the obstacles involved. We will focus on the case of unit capacities where $n = |A| = |O|$.

ii.1 Approach 1

First we prove the following corollary. Let $H[A, O]$ be a weighted bipartite graph in which the weight of every edge $(a, o) \in A \times O$, namely $w_{a,o}$, is iid from the uniform distribution over the interval $[0, 1]$.

Corollary ii.1 (Corollary of the Parisi Conjecture)**.** *Let δ_n denote the expected weight of the minimum weight matching in $H[A, O]$. Then,* $\lim_{n\to\infty} \delta_n \leq \frac{\pi^2}{6}$.

Proof sketch. The proof follows from the fact that the uniform distribution is stochastically dominated by the exponential distribution with mean 1, together with the Parisi conjecture. $\square$

Suppose that agent a prefers objects with lower weights in $H[A, O]$. Since each agent assigns n ranks $1, \ldots, n$ to the objects uniformly at random, then, due to the above corollary, one might conjecture that the expected rank of the rank-optimal assignment should be bounded by $\frac{\pi^2}{6}n$ (and thus the average rank by $\frac{\pi^2}{6}$). The next lemma establishes a relevant fact for a single agent.

Lemma ii.2. *Consider an edge $(a, o) \in A \times O$. The expected rank of o on agent a's preference list, namely $r_{a,o}$, equals $1 + \frac{n-1}{n} w_{a,o}$.*

Proof. Let $\gamma = w_{a,o}$. Object o has rank $r+1$ on agent a's list if there are precisely r edges adjacent to a that have a weight smaller than γ. We therefore can write

$$
\begin{aligned}
r_{a,o} &= \sum_{r=0}^{n-1} (r+1) \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} \\
&= \sum_{r=0}^{n-1} r \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} + \sum_{r=0}^{n-1} \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} \\
&= \sum_{r=1}^{n-1} (n-1) \binom{n-2}{r-1} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} + \sum_{r=0}^{n-1} \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} \qquad \text{(ii.1)} \\
&= \sum_{q=0}^{n-2} (n-1) \binom{n-2}{q} \left(\frac{\gamma}{n}\right)^{q+1} \left(1 - \frac{\gamma}{n}\right)^{n-2-q} + \sum_{r=0}^{n-1} \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} \\
&= \frac{\gamma}{n}(n-1) \sum_{q=0}^{n-2} \binom{n-2}{q} \left(\frac{\gamma}{n}\right)^{q} \left(1 - \frac{\gamma}{n}\right)^{n-2-q} + \sum_{r=0}^{n-1} \binom{n-1}{r} \left(\frac{\gamma}{n}\right)^r \left(1 - \frac{\gamma}{n}\right)^{n-1-r} \\
&= \frac{\gamma}{n}(n-1)(\frac{\gamma}{n} + (1 - \frac{\gamma}{n}))^{n-2} + (\frac{\gamma}{n} + (1 - \frac{\gamma}{n}))^{n-1} = \frac{\gamma}{n}(n-1) + 1, \qquad \text{(ii.2)}
\end{aligned}
$$

where (ii.1) follows from the fact that $r\binom{n-1}{r} = (n-1)\binom{n-2}{r-1}$, and (ii.2) follows from the Binomial theorem. □

Consider an assignment μ, and define $w_\mu = \sum_{a \in A} w_{a,\mu(a)}$. Then, by Lemma ii.2,

$$
\mathbb{E}\left[\sum_{a \in A} r_{a,\mu(a)}\right] \leq n(1 + w_\mu),
$$

which means that the expected average rank of μ is at most $1 + w_\mu$. At first glance, it might seem that taking μ to be the rank-optimal assignment and applying Corollary ii.1 implies that the expected average rank of the rank-optimal assignment is bounded by $1 + \frac{\pi^2}{6}$. This implication, however, is incorrect, because the above equality says nothing about the value of the conditional expectation

$$
\mathbb{E}\left[\sum_{a \in A} r_{a,\mu(a)} \middle| \mu \text{ is rank-optimal}\right].
$$

Computing this expression seems to be complicated. We have not found a way to compute it or to relate its value to the unconditional expectation preceding it.

ii.2 Approach 2

We next discuss a different approach that requires computing an alternative conditional expectation. We observe that

$$\mathbb{E}\left[\sum_{a\in A} w_{a,\mu(a)}\middle|\mu \text{ is rank-optimal}\right] = \sum_{a\in A}\mathbb{E}\left[w_{a,\mu(a)}\middle|\mu \text{ is rank-optimal}\right] = \sum_{a\in A}\frac{\mathbb{E}\left[r_a^*\right]}{n+1},$$

where r_a^* denotes the rank of agent a in the rank-optimal assignment. The last equality follows from the fact that the expected value of the k-th order statistic of n independent draws from the uniform distribution equals $\frac{k}{n+1}$ for $k \geq 1$. Therefore, to bound from above the expected average rank of the rank-optimal assignment, it suffices to bound

$$\mathbb{E}\left[\sum_{a\in A} w_{a,\mu(a)}\middle|\mu \text{ is rank-optimal}\right]$$

from above. Note, however, that Corollary ii.1 only provides a lower bound for this expression, because that corollary bounds the weight of the *minimum-weight* matching from above. (The weight of the minimum-weight matching is a lower bound on the weight of the rank-optimal matching.) Our Theorem 3.1 implies that this conditional expectation is bounded from above by a constant independent of n. An alternative method for bounding this expression, possibly using Corollary ii.1, is not known to us.