

Counting Houses of Pareto Optimal Matchings in the House Allocation Problem

Andrei Asinowski^{1*}, Balázs Keszegh^{2**}, and Tillmann Miltzow¹

¹ Institut für Informatik, Freie Universität Berlin, Germany

² Alfréd Rényi Institute of Mathematics, Budapest, Hungary

asinowski@gmail.com keszegh.balazs@renyi.mta.hu t.m@fu-berlin.de

Abstract. In an instance of the house allocation problem two sets A and B are given. The set A is referred to as *applicants* and the set B is referred to as *houses*. We denote by m and n the size of A and B respectively. In the house allocation problem, we assume that every *applicant* $a \in A$ has a preference list over every *house* $b \in B$. We call an injective mapping τ from A to B a *matching*. A *blocking coalition* of τ is a subset A' of A such that there exists a matching τ' that differs from τ only on elements of A' , and every element of A' improves in τ' , compared to τ according to its preference list. If there exists no blocking coalition, we call the matching τ a *Pareto optimal matching* (POM). A house $b \in B$ is *reachable* if there exists a Pareto optimal matching using b . The set of all reachable houses is denoted by E^* . We show

$$|E^*| \leq \sum_{i=1, \dots, m} \left\lfloor \frac{m}{i} \right\rfloor = \Theta(m \log m).$$

This is asymptotically tight. A set $E \subseteq B$ is *reachable* (respectively *exactly reachable*) if there exists a Pareto optimal matching τ whose image contains E as a subset (respectively equals E). We give bounds for the number of exactly reachable sets. We find that our results hold in the more general setting of multi-matchings, when each applicant a of A is matched with ℓ_a elements of B instead of just one. Further, we give complexity results and algorithms for corresponding algorithmic questions. Finally, we characterize *unavoidable* houses, i.e., houses that are used by all POM's. This yields efficient algorithms to determine all unavoidable elements.

1 Introduction

1.1 Definitions

In an instance of the house allocation problem two sets A and B are given. The set A is referred to as *applicants* and the set B is referred to as *houses*. We denote

* Research supported by the ESF EUROCORES programme EuroGIGA, CRP ‘Com-PoSe’, Deutsche Forschungsgemeinschaft (DFG), grant FE 340/9-1

** Research supported by Hungarian National Science Fund (OTKA), under grant PD 108406 and under grant NN 102029 (EUROGIGA project GraDR 10-EuroGIGA-OP-003) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the DAAD.

by m and n the size of A and B respectively. In the house allocation problem, we assume that every *applicant* $a \in A$ has a preference list over every *house* $b \in B$. We call an injective mapping τ from A to B a matching. A *blocking coalition* of τ is a subset A' of A such that there exists a matching τ' that differs from τ only on elements of A' , and every element of A' improves in τ' , compared to τ according to its preference list. If there exists no blocking coalition, we call the matching τ a *Pareto optimal matching* (POM).

The *underlying graph* is a complete bipartite graph on the set $A \cup B$. In this graph injective mappings indeed correspond to matchings.

We represent the preference lists by an $m \times n$ matrix. Every row represents the preference list of one of the applicants in A , i.e., in a given row r corresponding to some applicant $a \in A$, the leftmost house is the one that a prefers most, etc., house b_1 is left to b_2 in r if and only if a prefers b_1 over b_2 . Note that no row contains an element from B twice. We usually denote this matrix by M and following this interpretation we usually denote the applicants of A by r_1, r_2, \dots, r_m and the houses of B by $1, 2, \dots, n$. Because of this matrix representation, we usually refer to applicants of A only as rows and to houses of B as elements (of the matrix).

To illustrate the notion consider the following matrix and observe that the matching indicated by circles is indeed Pareto optimal.

$$\begin{pmatrix} \textcircled{1} & 5 & 3 & 2 & 4 \\ 3 & 1 & \textcircled{4} & 5 & 2 \\ 1 & \textcircled{3} & 5 & 4 & 2 \end{pmatrix}$$

With this notation the edge set of a matching τ in the underlying graph corresponds to a set of positions in the matrix. To make this formal consider an edge (a, b) in the underlying graph. Let r be the row of a in M and b the k^{th} house in row a . Then edge (a, b) corresponds to position $p = (r, k)$ in M . The image set of τ corresponds to the set of houses of B in these positions. Thus, we say that τ *picks, selects, chooses, reaches, assigns* some position p of M (resp. some element b of B), if p is in τ (resp. b is in the image set of τ). Similarly, we say that a row a *picks, selects, chooses, reaches, assigns* a position P in row a (resp. b) if this holds for the matching τ under consideration.

In a POM the positions after the m -th column will never be assigned, because at least one of the previous m elements in that row is preferred and not assigned to any other element on A . Therefore it is sufficient to consider only $m \times m$ square matrices.

If some POM τ assigns p (resp. b), then it is a *reachable* position (resp. *reachable* house). More generally, a set $E \subseteq B$ is *reachable* if there exists a POM τ with $E \subseteq s(\tau)$. In this case we also say that τ reaches E . A set E with $|E| = m$ is *exactly reachable* if there exists a Pareto optimal matching τ with $E = s(\tau)$. An element b is *unavoidable* if it belongs to the set $s(\tau)$ for every Pareto optimal matching τ of M and *avoidable* otherwise. A set E is *avoidable* if there exists a POM τ with $s(\tau) \cap E = \emptyset$. Note that for a set $|E| = m$ it is exactly reachable if and only if $B \setminus E$ is avoidable. We will also study matrices with fewer than m

columns, precise definitions will be given in Subsection 1.4. As a rule of thumb, in this case preference lists are shorter and it can happen that some elements of A are not assigned.

1.2 Results

Enumerating reachable elements and sets In Section 2 we deal with enumerative problems related to reachable elements. Our main result here is the following. Let M be an $m \times m$ matrix and E^* be the set of all reachable elements. Then

$$|E^*| \leq \sum_{i=1}^m \lfloor m/i \rfloor \leq m(\ln(m) + 1).$$

This improves the trivial upper bound of m^2 which appears in [8]. In [8] the authors also showed a lower bound construction which has asymptotically as many reachable elements as is implied by our upper bound. Thus Theorem 1.2 is asymptotically tight.

Denote by $\mathcal{E}(M)$ the family of all (exactly) reachable m -element sets of M . For example, if all the elements in the first column of M are distinct (or, more generally, if $|B| = m$), then $|\mathcal{E}(M)| = 1$. With Theorem 1.2 we can bound $\mathcal{E}(M)$.

Corollary 1. *For any matrix M , we have $|\mathcal{E}(M)| \leq \binom{m(\ln m + 1)}{m}$.*

This is the only non-trivial upper bound that we found, improving $\binom{m^2}{m}$ of [8]. As an important consequence, our upper bound also improves the upper bound on the pattern matching problem regarded in [8]. The best known lower bound is $\binom{m}{\lfloor m/2 \rfloor}$ [8]. The construction in that paper is a matrix where in the first $\lfloor m/2 \rfloor$ columns the i -th column c_i contains only element i and in the $(\lfloor m/2 \rfloor + 1)$ -st column there are m different elements which are also all different from $1, 2, \dots, \lfloor m/2 \rfloor$.

Characterization of avoidable elements and sets Section 3 concentrates on the notion of avoidable elements. Let x be the element suspect to be avoidable. Given some set of rows R we denote by $E_x(R)$ the set of elements left of x in the rows R (i.e., y is in $E_x(R)$ if and only if there exists a row $r \in R$ in which y appears to the left of x ; if x does not appear in R then all elements in R are regarded to be left of x).

An element x of a matrix M is avoidable if and only if for every set R of rows of M , we have:

$$|E_x(R)| \geq |R|$$

Extremal results and algorithmic results in connection to avoidable elements are included in Section 3.

Complexity of reachability Computational questions about reachable elements are considered detailed proofs can be found in the full version. We considered all reasonable computational questions connected to the notions we considered. The problems are defined as follows:

Problem 1 (Deciding Reachability)

Input: A matrix M and a set $D \subseteq B$.

Question: Is D reachable?

Problem 2 (Counting Exactly Reachable Supersets)

Input: A matrix M and some set $D \subseteq B$.

Question: How many sets E with $D \subseteq E \subseteq B$ are exactly reachable?

Problem 3 (Deciding Exact Reachability)

Input: A matrix M and a set $E \subseteq B$, $|E| = m$.

Question: Is E (exactly) reachable?

Problem 4 (Counting Reachable Sets)

Input: A matrix M .

Question: How many sets $D \subseteq B$ are reachable?

Problem 5 (Counting Exactly Reachable Sets)

Input: A matrix M .

Question: How many sets E are exactly reachable?

The next table summarizes our findings about algorithmic questions. The general case is always the same as with 3 column matrices. Problems 1 and 2 are already complete if D contains exactly 1 element. Our contribution among others is to show NP-completeness also for matrices with only 3 columns.

Problem	2 columns	3 columns
1) Deciding Reachability	polynomial	NP-complete
2) Counting Exactly Reachable Supersets	#P-complete	#P-complete
3) Deciding Exactly Reachability	polynomial	polynomial
4) Counting Reachable Sets	explicit formula	?
5) Counting Exactly Reachable Sets	#P-complete	#P-complete

It remains an open question whether Problem 4 is hard for general matrices. We conjecture it is already #P-complete for 3 column matrices.

1.3 Motivation and related work

One-sided matchings have natural practical uses, e.g. consider the house-allocation problem where the set A consists of people and the set B consists of houses, see for instance [2].

A recent book on matchings under preferences is by David Manlove [9]. In this paper we tried, whenever applicable, to follow the notation therein.

A field that evidently seems to be related to our topic is that of stable matchings. This field is very broad and belongs to economic game theory. The seminal

work from Gale and Shapley is the starting point for this field [7]. Some work in this field and different variations of the problem can be found in the PHD thesis of Sandy Scott [13], recent papers can be found in the online available proceedings of the Second International Workshop on Matching Under Preferences called MATCH UP [1]. In these works there are many different concepts of preferences and stability and they ask for efficient computable solutions that maximize the outcome for the participants in one way or the other. Readers interested more broadly in the topic of algorithmic game theory are referred to the book edited by Nisan, Roughgarden, Tardos and Vazirani [10].

In contrast to most research done in these areas, our question is more combinatorial in nature. The underlying algorithmic question of computing a Pareto optimal matching is trivial. Thus, instead of existence questions, rather the enumerative questions become interesting. However, for the original definition of stability many authors have tried to upper and lower bound the number of stable matchings and some combinatorial structures have been unfolded. See Section 2.2.2 [9] for an overview of results in this direction.

Further some of the complexity results we will present have been found, in parallel and without our awareness. The first dates back to 2005 [3]. Their main result is an efficient algorithm to compute a POM with maximal cardinality. Further they show hardness to compute a minimum maximal POM. The first results already has some ideas of the proof of Theorem 1.2. Although they show an easy 2-approximation, it is open, whether there exists a PTAS for a minimum maximal POM.

We are aware of 4 more papers that considered similar results to our complexity results [12], [4], [5] and [6]. All of them appeared in 2013 three of them in December. Their main motivation is to study the behavior of the randomized serial dictatorship also called randomized priority allocation. The randomized serial dictatorship picks a permutation at random and thereafter computes the corresponding greedy matching.

Saban and Sethuraman solved, in this context, NP-hardness of Problem 1, for arbitrary matrices [12]. Note that Henze, Jaume and Keszegh showed first that Problem 1 is NP-complete [8]. Aziz, Brandt and Brill showed #P-hardness for a variant of Problem 2 for arbitrary matrices [4]. We improve these results, as we can show this holds also for matrices with only 3 columns. Aziz and Meske show that constraint versions are solvable in polynomial time [5]. At last Cechlárová et. al. consider a generalized setting. However they show NP-hardness of compute a minimum maximal matching even for matrices with 2 columns by an elegant and simple reduction from vertex-cover [6].

Another important connection is that this work is originally motivated by a work that was presented at the EuroCG 2012 in Braunschweig [8]. The authors considered a generalisation of Voronoi diagrams under the assumption that not just one point, but many points are matched injectively to a ‘nearest neighbor’, in a way that minimizes the sum of the square root of distances between matched points. From the definitions in their paper, the Pareto optimality comes as a natural property. They asked explicitly for the number of exactly reachable sets,

as it gives an upper bound on the number of Voronoi cells in the above setting. Motivated by this, they gave lower and upper bounds on the number of exactly reachable stable sets. To do this, first they gave lower and upper bounds for the number of reachable elements. In this paper we improve their upper bound for the number of reachable elements and by that we prove that their lower bound is asymptotically correct. This also yields a significant improvement on the previous upper bound on the number of exactly reachable stable sets, although in this case our new upper bound still does not meet the lower bound they had.

Their work is based on a work by Rote presented at the EuroCG 2010 (2 years earlier) in Dortmund [11].

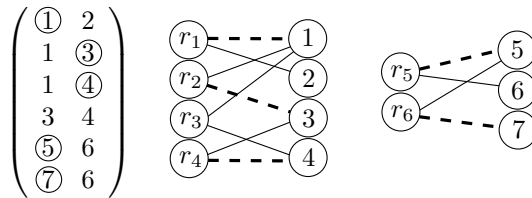
1.4 Preliminaries

As we also want to study matrices with fewer than m columns, we need to define what we mean by a matching under these assumptions. There are two equivalent ways. First we could say that every row, for which all elements are already picked by other rows just do not get assigned to anything. A nicer way is to add columns, with all elements in one column being the same and not appearing before. If we want to know if some set E is exactly reachable in the first way, we construct E' from E by adding the elements from the first $m - |E|$ additional columns (and vice versa). The following is an example of a 2 column matrix.

$$\begin{pmatrix} 1 & 4 \\ 2 & 1 \\ 2 & 5 \\ 4 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & c_1 & c_2 \\ 2 & 1 & c_1 & c_2 \\ 2 & 5 & c_1 & c_2 \\ 4 & 3 & c_1 & c_2 \end{pmatrix}$$

We use the first approach. However, using the second approach, some hardness results will carry over from 2 or 3 column matrices to k column matrices ($2 \leq k \leq m$). In such a case, we will point this out again at the appropriate places.

To see the correspondence between matchings in a graph theoretical sense and in our context we define the *bipartite row element graph* G as follows. The vertices are defined as the set of rows and elements; an element e is adjacent to some row r if and only if e appears in r . See an example for the special case of a matrix with only 2 columns.



The circled POM corresponds to the dashed matching on the right side.

If there is no blocking coalition of size $\leq i$, we call the matching an *i-Pareto optimal matching* (*i-POM*). In particular this implies that every POM is an *i-POM*. We call a matching *1-POM* if there is no blocking coalition of size one. The next matching is one-Pareto optimal but not Pareto optimal.

$$\begin{pmatrix} 1 & \textcircled{5} & 3 \\ 5 & 1 & \textcircled{4} \\ 5 & \textcircled{1} & 1 \end{pmatrix}$$

A matching τ is *greedy* if there exists a permutation π of A such that the matching can be generated in the following manner: we process the rows of M in the order determined by π , and in each row we pick the leftmost element that was not picked earlier. Given some permutation π we call the corresponding greedy matching τ_π .

Lemma 1 brings all the introduced notions together, showing that POM, 1-POM and greedy matchings select exactly the same sets. The equivalence of POM and greedy matchings was already proved in [8].

Lemma 1. *Let $E \subseteq [n]$ with $|E| = m$. The following statements are equivalent.*

1. *E is (exactly) reachable, i.e. there exists a POM τ with $s(\tau) = E$.*
2. *There exists a permutation π such that for the greedy matching τ_π we have $s(\tau_\pi) = E$.*
3. *There exists an one-Pareto optimal matching (1-POM) τ with $s(\tau) = E$.*

Proof. [1 \Rightarrow 2] Let τ be a POM matching such that $s(\tau) = E$. We construct a permutation π inductively. If possible take as the next row, in the order of our permutation, the one that has a position of τ in its first entry. Delete the element a at this position from all other rows and continue. We show that at each stage there must be such a row. For the purpose of contradiction assume such a row does not exist. Take any row, denoted by q_1 and let e_1 be some element which is left to the element selected by τ in row q_1 . Because τ is Pareto optimal, there exists some row r_2 selecting e_1 . Let e_2 be any element left to e_1 in row r_2 . In this way we can define a sequence (e_i) and (r_i) . As we have only finitely many elements, at some point we get a first e_j that appears earlier in the sequence $e_i = e_j$, $i < j$. This implies that in the rows r_i, \dots, r_j we can improve simultaneously (i.e., it is a blocking coalition), which is a contradiction to the assumption that τ is Pareto optimal.

[2 \Rightarrow 3] As every row picks the best element, not yet selected, it is clear that no single row can improve.

[3 \Rightarrow 1] Let τ_0 be some 1-Pareto optimal matching and $E = s(\tau_0)$. Observe that all the elements left to the elements picked by τ are in E . The set of matchings that are better or equal to τ_0 is non-empty as it contains τ_0 and the set is of course finite, so there exists a best matching τ_1 among them, i.e. one for which there is no better matching. This must be a POM and by Observation 2 $s(\tau_1) = s(\tau_0) = E$, and the size of $s(\tau_1)$ is also m . \square

Note that this lemma implies that also for any i , i -POMs select the same sets as POMs/1-POMs. Note also that the proof of Lemma 1 implies that actually every greedy matching is Pareto optimal and vice versa.

2 Enumerating reachable elements and sets

We start with a trivial but important observation.

Lemma 2. *If τ is a POM and τ selects position p in row a , then τ selects every element that appears in row a left of p .*

For every row r , there exists a reachable position p_r furthest to the right in that row, we call such a position *last reachable*. However note, that not all positions must be reachable left of the last reachable position. Consider the following matrix. Together with the matching τ indicated by circles.

$$\begin{pmatrix} \textcircled{5} & 4 & 3 & 2 \\ 5 & \textcircled{1} & 6 & 7 \\ 1 & \textcircled{2} & 8 & 9 \\ 2 & 1 & 5 & \textcircled{4} \end{pmatrix}$$

Clearly, τ is a POM and thus the circled position in the bottom row with element 4 is the last reachable position in that row. However, it is easy to check, that the two positions left to this circled position (with elements 1 and 5) are not reachable.

Let M be an $m \times m$ matrix and E^* be the set of all reachable elements. Then

$$|E^*| \leq \sum_{i=1}^m \lfloor m/i \rfloor \leq m(\ln(m) + 1).$$

Proof. Let τ_i be a POM selecting the last reachable position p_i in row i ($1 \leq i \leq m$) (these matchings are not necessarily different.).

Let e be some element that can be reached by some POM. We show e is selected by one of the POMs τ_1, \dots, τ_m . Indeed, if e is at some last reachable position then this is clear. Otherwise, e appears in some row r not at the last position p_r . By Observation 2, e must be picked by τ_r . Thus the matchings $\tau_1 \dots, \tau_m$ reach together all reachable elements. As $\tau_1 \dots, \tau_m$ are m POMs, the first inequality follows from Lemma 3. Finally, it is well-known that the harmonic series is bounded by $\ln(m) + 1$, thus the second inequality holds as well. \square

Lemma 3. *Let T be some set of k POMs. We denote by $E(T)$ the set of elements reached by at least one POM of T . Then*

$$|E(T)| \leq \sum_{i=1}^k \lfloor m/i \rfloor.$$

Proof. The proof goes by induction on k . The base case $k = 1$ is true as one POM selects exactly m different elements.

Consider now a set T of $k \geq 2$ POMs and the set of positions reached by T . Among these positions we denote by p_i the position furthest to the right in row i and we denote $F = \{p_1, \dots, p_m\}$. We say that an element e (resp. position p) is uniquely reachable by some τ if τ is the only POM in T that reaches e (resp. selects p). Consider the set $G \subseteq F$ of those rightmost reachable positions that are reachable by exactly one POM of T . By the pigeon-hole principle there exists a POM τ in T that reaches at most $1/k$ portion of G . Denote the set of elements in these positions by H ($|H| \leq \lfloor m/k \rfloor$).

By the definition of H all other elements are not selected uniquely by τ , i.e. some other matching of T also selects it. Thus the rest of the reached elements are also reachable by $T - \tau$. By induction we get

$$E(T) \leq E(T - \tau) + \lfloor m/k \rfloor \leq \left(\sum_{i=1}^{k-1} \lfloor m/i \rfloor \right) + \lfloor m/k \rfloor = \sum_{i=1}^k \lfloor m/i \rfloor.$$

This finishes the proof. \square

Next we show two constructions concerning tightness of the results from Theorem 1.2 and Lemma 3.

Asymptotic tightness of Theorem 1.2 follows from the following construction by Henze, Jaume and Keszegh [8].

Example 1 ([8]). For each k , a matrix M_k with $m = 2^k$ rows and $(m/2) \log 4m = (k+2)2^{k-1}$ reachable elements is constructed recursively as follows.

$$M_0 = (1);$$

and, for $k \geq 0$,

$$M_{k+1} = \left(\begin{array}{c|c} 1 & \\ \vdots & M'_k \\ \hline 2^k & \\ \hline 1 & \\ \vdots & M''_k \\ 2^k & \end{array} \right),$$

where M'_k and M''_k are relabelings³ of M_k with no common element and all elements different from $1, 2, \dots, 2^k$. The undefined entries of the matrix can be filled arbitrarily.

Regarding Lemma 3, we prove that it is tight for certain values of k and m :

³ A matrix M' is a *relabeling* of a matrix M if there is a bijective function between the elements (not positions!) of M and M' such that applying this function to the elements in all the positions of M we get M' . Clearly two matrices that are relabelings of each other are equivalent from our perspective.

Corollary 2. For every k there exists a matrix N_k with $m = k!$ rows and a set T_k of k POMs, such that the number of elements reached by T_k is exactly $\sum_{i=1}^k m/i$.

Proof. The construction is again recursive. For each k we define the matrix N_k with $m = k!$ rows and k columns with the property that each element appears only in one column, and each element that appears in the j th column ($j \leq k$), appears there exactly $k-j+1$ times. We also define a set Π_k of k permutations of the $k!$ rows from which we get T_k by taking the greedy matchings corresponding to the permutations. We will prove that all the elements of N_k are reachable by some greedy matching of T_k .

The matrices N_k are defined in the following way:

$$N_1 = (1),$$

and for $k \geq 1$:

$$N_{k+1} = \left(\begin{array}{c|c} 1 & \\ \vdots & N_k^1 \\ \hline k! & \\ \vdots & \\ \hline 1 & \\ \vdots & N_k^{k+1} \\ \hline k! & \end{array} \right).$$

Here $N_k^1, N_k^2, \dots, N_k^{k+1}$ are $k+1$ matrices which are all relabelings of N_k with no elements common to each other and to the set $\{1, 2, \dots, k!\}$. It is clear that N_{k+1} has $(k+1)!$ rows and $k+1$ columns. Moreover, each element in the j th column ($j \leq k+1$) appears there $(k+1) - j + 1$ times: this is clear for the first column, and is easily seen for other columns by induction.

Next we define the permutations. For $k = 1$, Π_1 contains the only permutation on the one row of N_1 . Next we recursively define Π_{k+1} . For each N_k^j ($1 \leq j \leq k+1$), we have by recursion an associated set $\{\pi_1^j, \pi_2^j, \dots, \pi_k^j\}$ of k permutations (thus, π_i^j is the i th permutation from Π_k relabeled accordingly to N_k^j – the j th copy of N_k in N_{k+1}). Now the permutations in Π_{k+1} are defined as follows. For every i ($1 \leq i \leq k+1$), the permutation π_i is obtained by taking first the $k!$ rows of N_k^i in any (for example, the natural) order; then the rows of $N_k^1 \cup N_k^2 \cup \dots \cup N_k^{i-1}$ in the order determined by the permutations $\pi_{i-1}^1, \pi_{i-1}^2, \dots, \pi_{i-1}^{i-1}$; and, finally, the rows of $N_k^{i+1} \cup N_k^{i+2} \cup \dots \cup N_k^{k+1}$ in the order determined by the permutations $\pi_i^{i+1}, \pi_i^{i+2}, \dots, \pi_i^{k+1}$. Clearly, each row was taken once, so π_i is indeed a permutation. Also, when processing such a permutation, in the first $k!$ steps we choose all elements $1, 2, \dots, k!$, so in the rest the permutation chooses the same elements in each N_k^j ($j \neq i$) as the corresponding permutation (π_{i-1}^j or π_i^j) would choose in N_k^j .

Thus by induction it is true that these permutations choose all elements of N_k . Indeed, this is true for N_1 and by induction it remains true as for every N_k^j .

($1 \leq j \leq k + 1$) all π_i^j ($1 \leq i \leq k$) is part of some π_u ($1 \leq u \leq k + 1$). Finally, the number of different elements in N_{k+1} is $\sum_{i=1}^{k+1} m/i$, as we have $k + 1$ columns and in the j th column ($1 \leq j \leq k + 1$) each element appears $(k + 1) - j + 1$ times, thus this column has $\frac{m}{(k+1)-j+1}$ different elements. \square

3 Characterization of avoidable elements

In this section we give characterization of avoidable elements. Recall that we define $E_x(R)$ as the set of elements left of x in the rows of R (i.e., y is in $E_x(R)$ if and only if there exists a row $r \in R$ in which y appears to the left of x ; if x does not appear in R then all elements in R are regarded to be left of x)

An element x of a matrix M is avoidable if and only if for every set R of rows of M , we have:

$$|E_x(R)| \geq |R|$$

Proof. $[\Rightarrow]$ Let τ be a POM which does not pick x and let R be a set of rows. In each row a different element is picked by τ , which is left of x . This shows the claim.

$[\Leftarrow]$ W.l.o.g. x is present in all the rows. Consider the bipartite graph on $A \cup B$, defined by all pairs $(a, b) \in A \times B$ such that b appears in row a before x . The above condition says, that for all subsets $R \subset A$ the neighbourhood of R is larger or equal to R in terms of size.

By Hall's theorem, there exists a matching τ that picks elements to the left of x . W.l.o.g. in τ each row picks an element farthest to the left in M not chosen by any other row. In other words τ is an 1-POM. By Lemma 1 there is a POM τ' selecting the same set of elements as τ , thus τ' does not choose x and so x is avoidable.

Acknowledgments We want to thank Matthias Henze and Rafel Jaume for posing this open question. We also want to thank Rob Irving, Ágnes Cseh and David Manlove for helping us to find related work to our problem. Special thanks goes to Nieke Aerts for enjoyable and interesting discussions on attempts to improve Corollary 1.

This research was partially done while the second author was at FU Berlin in the scope of an EuroGIGA Cross-CRP visit and later with a DAAD Study Visit Grant for Senior Academics.

References

1. *MATCH-UP 2012: the Second International Workshop on Matching Under Preferences*, Corvinus University of Budapest, Hungary, 2012.
2. Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

3. David J Abraham, Katarína Cechlárová, David F Manlove, and Kurt Mehlhorn. Pareto optimality in house allocation problems. In *Algorithms and Computation*, pages 1163–1175. Springer, 2005.
4. Haris Aziz, Felix Brandt, and Markus Brill. The computational complexity of random serial dictatorship. *Economics Letters*, 121(3):341–345, 2013.
5. Haris Aziz and Julian Mestre. Parametrized algorithms for random serial dictatorship. *arXiv preprint arXiv:1403.0974*, 2014.
6. K Cechlárová, P Eirínakis, T Fleiner, D Magos, I Mourtos, and E Potpinková. Pareto optimality in many-to-many matching problems. *preprint*, 2013.
7. David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.
8. Matthias Henze, Rafel Jaume, and Balsz Keszegh. On the complexity of the partial least-squares matching voronoi diagram. In *Proceedings of the 29th European Workshop on Computational Geometry (EuroCG)*, pages 193–196, March 2013.
9. David Manlove. *Algorithmics of matching under preferences*. World Scientific Publishing, 2013.
10. Noam Nisan. *Algorithmic game theory*. Cambridge University Press, 2007.
11. Günter Rote. Partial least-squares point matching under translations. In *26th European Workshop on Computational Geometry (EuroCG10)*, pages 249–251, March 2010.
12. Daniela Saban and Jay Sethuraman. The complexity of computing the random priority allocation matrix. In *Web and Internet Economics*, pages 421–421. Springer Berlin Heidelberg, 2013.
13. Sandy Scott. *A study of stable marriage problems with ties*. PhD thesis, University of Glasgow, 2005.