

Capacity Variation in the Many-to-one Stable Matching[★]

To Gerhard Woeginger (1964-2022), an outstanding computer scientist

Federico Bobbio^{1a,b}, Margarida Carvalho^a, Andrea Lodi^d, Alfredo Torricco^c

^aCIRRELT and Département d'informatique et de recherche opérationnelle, Université de Montréal

^bCERC in Data Science, Polytechnique Montréal

^cCornell University

^dJacobs Technion-Cornell Institute, Cornell Tech

Abstract

We study the computational complexity of optimizing hospital capacities in the many-to-one stable matching problem, a key model for markets like hospital-resident assignments. Unlike the classical fixed-capacity setting, we allow limited capacity changes to improve resident outcomes while maintaining stability and respecting a budget. This variant has real-world relevance but unclear theoretical complexity. We prove that deciding the optimal capacity expansion is NP-complete and the optimization problem is inapproximable within $n^{\frac{1}{6}-\epsilon}$, where n is the number of hospitals.

1. Introduction

The stable matching problem in a bipartite graph is a classical problem that requires a few key ingredients to be defined: Two disjoint sets of nodes, a set of edges connecting the nodes in the two sets, a capacity function for each node and a preference ranking of the node over the edges. A matching is a collection of edges that respects the capacity of each node, and it is said to be stable when such collection cannot be simultaneously improved for two nodes of every connected pair in the graph. It has been shown that a stable matching always exists via a polynomial-time algorithm [10]. If we assume the graph to be complete and we assume some edges are tied in the ranking of the node, then the problem of finding a maximum cardinality stable matching remains polynomial. Nonetheless, as soon as we relax also the assumption that the bipartite graph is complete, then the problem of finding a stable matching of maximum cardinality becomes NP-hard [15].

The stable matching problem has found multiple real-world applications. These include admission to education systems [19, 4], the integration of refugees in hosting countries [9], and the rationing of healthcare services such as ventilators during a pandemic [3]. Most of these applications assume that one of the sets of nodes (the hospitals) can have the capacity function greater than one, while the other (the residents) has capacity one. This is known as the many-to-one stable matching (HR) problem, where HR stands for hospital-resident. Henceforth, we refer to one set of nodes in the graph as hospitals and

the other as residents.

In the standard version of HR, the capacity of each hospital is fixed, and the decision-maker selecting a matching (assignment) has no control over these quotas. However, in many real-life situations, variations in market size—either expansion or reduction—can play a significant role. For example, the University of Tsukuba recently restructured its course offerings by allocating a budget of capacities to courses [14]. In Bobbio et al. [5, 6], mathematical programming tools are introduced to optimize the allocation of extra capacities, particularly within the Chilean school choice system. This highlights the practical importance of considering the HR with capacity variation. Specifically, there is a need to produce solutions for this problem version, *i.e.*, feasible assignments of hospital capacities and associated stable matchings, optimizing a specific objective.

In what follows, we first establish the notation for the many-to-one stable matching problem. Subsequently, we provide our main problem definition, which centers on expanding hospital capacities with the goal of finding a stable matching that benefits the residents under a given budget of extra capacity.

Notation.. Let $\mathcal{R} = \{i_1, \dots, i_{|\mathcal{R}|}\}$ be the set of residents, and let $\mathcal{H} = \{j_1, \dots, j_{|\mathcal{H}|}\}$ be the set of hospitals, with edges $\mathcal{E}$ between residents and hospitals whenever they find each other acceptable. We assume here that every resident-hospital pair is acceptable, *i.e.*, $\mathcal{E} = \mathcal{R} \times \mathcal{H}$. Each hospital $j \in \mathcal{H}$ has a non-negative integer capacity c_j , representing the maximum number of residents it can admit. Formally, a matching M is a subset of $\mathcal{E}$ in which each resident is matched to at most one hospital, and each hospital j has at most c_j residents assigned in M .

An instance of the many-to-one stable matching problem is a tuple $\Gamma = \langle \mathcal{R}, \mathcal{H}, \succ, \mathbf{c} \rangle$, where $\succ$ denotes the preference lists for both the residents and the hospitals, and $\mathbf{c}$ is the vector of hospital capacities. For a resident i , $j \succ_i j'$ means i prefers hospital j over j' . Hospital preferences are assumed to be responsive: each hospital's ranking is induced by its linear order

[★]Accepted manuscript. The Version of Record is available at <https://doi.org/10.1016/j.orl.2026.107416>. © 2026. This manuscript version is made available under the CC BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Email addresses: federico.bobbio@umontreal.ca (Federico Bobbio), carvalho@iro.umontreal.ca (Margarida Carvalho), andrea.lodi@cornell.edu (Andrea Lodi), alfredo.torricco@cornell.edu (Alfredo Torricco)

over individual residents. For a matching M , we say that (i, j) is a blocking pair if (i) resident i is unmatched or prefers j to her assigned hospital in M , and (ii) j is not at full capacity or prefers i to at least one resident it currently has in M . The matching M is stable if it admits no blocking pair.

Whenever all preference lists are strict and complete, the well-known deferred acceptance (DA) algorithm [10] finds a unique resident-optimal stable matching, meaning that no resident can get a strictly better hospital in any other stable matching. Finally, given $n \in \mathbb{Z}$, we define $[n] = \{1, \dots, n\}$.

Objective.. We measure the quality of a matching via the average (or total) rank of the assigned hospitals from the perspective of residents. Formally, if $\text{rank}_i(j)$ denotes the position of j in i 's preference list (with 1 being the top choice), then the un-normalised average hospital rank of a matching M is $\text{AvgRank}(M) = \sum_{(i,j) \in M} \text{rank}_i(j)$. This objective is consistent with finding a stable matching that is most preferred by the residents (i.e., “resident-optimal”), as minimizing this sum identifies the stable matching that yields the best overall outcome for residents.

1.1. Problem Definition

In many applications, there is a budget of extra hospital capacities that a planner can distribute among the hospitals. Formally, let $\mathbf{t} \in \mathbb{Z}_+^H$ be a non-negative integer vector denoting how many extra seats are allocated to each hospital. Then, the instance $\Gamma_{\mathbf{t}} = \langle \mathcal{R}, \mathcal{H}, >, \mathbf{c} + \mathbf{t} \rangle$ has hospital capacities $c_j + t_j$ for each $j \in \mathcal{H}$. Suppose further that there is a budget $B \in \mathbb{Z}_+$ such that we require $\sum_{j \in \mathcal{H}} t_j = B$. We are interested in whether there exists a capacity allocation $\mathbf{t}$ of size B and a corresponding stable matching of $\Gamma_{\mathbf{t}}$ whose average rank is below some target K . Formally:

Problem 1 (MIN-AVG_{EXP} HR).

INSTANCE: A HR instance $\Gamma = \langle \mathcal{R}, \mathcal{H}, >, \mathbf{c} \rangle$, a non-negative integer $B \in \mathbb{Z}_+$ (extra seats), and a target value $K \in \mathbb{Z}_+$.

QUESTION: Does there exist a non-negative vector $\mathbf{t} \in \mathbb{Z}_+^H$ such that $\sum_{j \in \mathcal{H}} t_j = B$ and a stable matching $M_{\mathbf{t}}$ in $\Gamma_{\mathbf{t}}$ (with capacities $\mathbf{c} + \mathbf{t}$) for which $\text{AvgRank}(M_{\mathbf{t}}) \leq K$?

As we discuss in the sequel, finding such a capacity allocation and stable matching is computationally challenging: we show that MIN-AVG_{EXP} HR is NP-complete and its optimization version cannot be approximated within a factor of $(n)^{\frac{1}{6}-\varepsilon}$ (where n is the number of hospitals), unless $P = NP$. Our complexity proof is based on a new structure that we call *village*. Each village is assigned some extra capacities, and the preferences of the hospitals and residents in a village ensure that the extra capacities can be optimally allocated only in a specific way. The role of the village is to reproduce the functioning of ties in a stable matching where agents have strict preferences. The structure of the village can be also used to prove a similar inapproximability result for the problem of reducing capacities; the capacity reduction problem and the problem of capacity variation in a partition of the set of hospitals are introduced and studied in an expanded version of this work [7].

2. Related Work

General context. In their seminal paper, Gale and Shapley [10] introduce the stable matching problem and provide a polynomial-time algorithm, i.e., DA. In general, the main focus of the literature has been on finding the maximum cardinality stable matching, which can be efficiently obtained when there are incomplete preference lists without ties or complete preference lists that include ties. Once we assume both incomplete lists and ties, finding the maximum cardinality stable matching becomes NP-hard [15].

Capacity variation. In economics, the design of a stable matching mechanism, when the number of participants of one side is increased, is known as the *entry comparative static*. In [12, 11, 13], the authors prove that when a new hospital is added to the instance, all residents are matched *weakly better*. A substantial part of the literature has focused on strategy-proof matching mechanisms, i.e., on matching mechanisms that incentivize participants to reveal their true preferences. Sönmez [20] proves that hospitals can manipulate the stable matching in their favor by falsely reporting a reduced capacity. Moreover, Romm [18] proves that the stable matching mechanism can still be manipulated even if the reported capacities are used during the admission process. Yahiro and Yokoo [21] consider a profile of “resources” that can be allocated to “projects” (hospitals) and focus on designing strategy-proof and efficient mechanisms. Nguyen and Vohra [16] study the problem of ensuring stability in a matching market with couples, and find that by adding at most 9 extra capacities, the existence of a stable matching is guaranteed. Contrary to the approach of considering extra capacities as a parameter, contemplated in the entry comparative static, Bobbio et al. [5, 6] consider for the first time the allocation of extra capacities to hospitals as a decision variable rather than a parameter. A heuristic solution methodology to optimize the outcome for the residents is devised in [1]. Recently, [2] also studies the problem of allocating capacities, the authors propose a general framework to devise a mechanism and they study its game theoretic properties. Kumano et al. [14] study the problem from both a theoretical and practical perspective, their work was used to guide the reallocation of quotas at the University of Tsukuba in Japan. Chen and Csáji [8] study the computational challenges of capacity augmentation, with the goal of minimizing the maximum capacity increase or minimizing the sum of capacity increase. Finally, Ranjan et al. [17] prove that achieving *strongly stable* matchings often necessitates quota augmentation, and provide a polynomial-time solution for minimizing total quota increases.

3. The Capacity Expansion Problem

Our main result establishes the computational complexity and inapproximability of Problem 1. Denote by MIN-AVG_{EXP} HR OPT the optimization version of Problem 1.

Theorem 3.1. MIN-AVG_{EXP} HR is NP-complete. Moreover, unless $P=NP$, for any $\varepsilon > 0$, MIN-AVG_{EXP} HR OPT cannot be approximated within a factor of $n^{(\frac{1}{6}-\varepsilon)}$, where n is the number of hospitals.

To prove Theorem 3.1, we do a reduction from the problem of minimizing the cost when there are *ties*. The stable marriage (SM) problem corresponds to the HR problem where $c_j = 1$ for all $j \in \mathcal{H}$. We call SMT the case in which there are ties in the SM problem. A tie, denoted by $(\cdot)$, appears when an agent has in the same position of the list two different participants of the opposite side. For example, if the preference list for resident i is $j_4, (j_1, j_3), j_2$, then the rankings of $j \in \{j_1, j_3\}$ are $\text{rank}_i(j) = 2$, and $\text{rank}_i(j_2) = 4$. In SMT, a matching M is *weakly stable* if there is no pair such that both agents *strictly* prefer each other over their allocation in M .

Problem 2 (MIN-w SMT).

INSTANCE: An SMT instance $\Gamma = \langle \mathcal{R}, \mathcal{H}, >, \mathbf{c} \rangle$ with $c_j = 1$ for all $j \in \mathcal{H}$ and a target value $K \in \mathbb{Z}_+$.

QUESTION: Is there a weakly stable matching M such that $\text{AvgRank}(M) \leq K$?

We use MIN-w SMT OPT to denote the optimization version of MIN-w SMT, i.e., the problem of finding a weakly stable matching that minimizes AvgRank. Building upon the ideas in [15], we can obtain the following result for MIN-w SMT.

Lemma 3.1. MIN-w SMT is NP-complete. Moreover, for any $\varepsilon > 0$, MIN-w SMT OPT is not approximable within a factor of $n^{1-\varepsilon}$, unless $P = NP$, where $n = |\mathcal{H}|$. This result holds even if ties are only on residents' side, there is at most one tie per list, and each tie is of length two.

The proof of Lemma 3.1 appears in Appendix Appendix A.1. The technique used there is necessary, but not sufficient, for proving Theorem 3.1: preserving the matching structure in the reduced instance, requires a careful allocation of the additional capacities. To address this, the next section introduces a new structure, the *village*.

3.1. Design of the Instance

We focus on instances of SMT satisfying the properties of Lemma 3.1, and we introduce a polynomial transformation to instances of MIN-AVG_{EXP} HR. Let $\Gamma = \langle \mathcal{R}, \mathcal{H}, >, \mathbf{c} \rangle$ be an instance of MIN-w SMT such that $|\mathcal{R}| = |\mathcal{H}| = n$. Let $L \leq n$ be the number of residents with ties in their preference list. The set of residents is partitioned in two sets $\mathcal{R} = \mathcal{R}' \cup \mathcal{R}''$, where $\mathcal{R}'$ is the set of residents with a tie of length two at the head of the preference list and $\mathcal{R}''$ is the set of residents with a strict preference list. Henceforth, we fix an ordering of the residents in $\mathcal{R}$ and denote $\mathcal{R}' = \{i_1, \dots, i_L\}$ and $\mathcal{R}'' = \{i_{L+1}, \dots, i_n\}$. Since preference lists are complete, observe that in any weakly stable matching every resident is matched. Next, we create an instance $\hat{\Gamma} = \langle \hat{\mathcal{R}}, \hat{\mathcal{H}}, \hat{>}, \hat{\mathbf{c}} \rangle$ of MIN-AVG_{EXP} HR with a specific target value and budget.

Hospitals and residents. First, we add n^4 copies of the residents in $\mathcal{R}''$, $\bar{\mathcal{R}}'' = \{i_{L+1}^k, \dots, i_n^k\}_{k \in [n^4]}$, and we introduce the set of residents $\mathcal{A} = \{a_1, \dots, a_{n^6}\}$. Moreover, we create n^4 copies of the hospitals in $\mathcal{H}$: $\mathcal{H} = \{j_1^k, \dots, j_n^k\}_{k \in [n^4]}$. We also introduce a set $\mathcal{Z} = \{z_1, \dots, z_{n^6}\}$ with n^6 hospitals, and a set $\mathcal{X} = \{x_1, \dots, x_{n^7}\}$ with n^7 hospitals.

For every resident $i_\ell \in \mathcal{R}'$ ($\ell \in [L]$), and $k \in [n^4]$ we introduce a structure $\mathcal{B}_\ell^k$ that we call *village*, which is composed of

- a set of residents $\mathcal{W}_\ell^k = \{w_{\ell,e}^k\}_{e \in [n]}$,
 - a resident y_ℓ^k , and
 - two sets of hospitals $\mathcal{V}_\ell^k = \{v_{\ell,e}^k\}_{e \in [n]}$ and $\bar{\mathcal{V}}_\ell^k = \{\bar{v}_{\ell,e}^k\}_{e \in [n]}$.
- Two sets of residents $\mathcal{U}_\ell^k = \{u_{\ell,e}^k\}_{e \in [n]}$ and $\bar{\mathcal{U}}_\ell^k = \{\bar{u}_{\ell,e}^k\}_{e \in [n]}$.

Generally, we denote as $\mathcal{V} := \bigcup_{\ell=1}^L \bigcup_{k=1}^{n^4} \mathcal{V}_\ell^k$, $\bar{\mathcal{V}} := \bigcup_{\ell=1}^L \bigcup_{k=1}^{n^4} \bar{\mathcal{V}}_\ell^k$, $\mathcal{U} := \bigcup_{\ell=1}^L \bigcup_{k=1}^{n^4} \mathcal{U}_\ell^k$ and $\bar{\mathcal{U}} := \bigcup_{\ell=1}^L \bigcup_{k=1}^{n^4} \bar{\mathcal{U}}_\ell^k$. We also define $\mathcal{Y} := \bigcup_{\ell=1}^L \bigcup_{k=1}^{n^4} \{y_\ell^k\}$, and $\mathcal{V}^k := \bigcup_{\ell=1}^L \mathcal{V}_\ell^k$. In summary, the new instance $\hat{\Gamma}$ is made of a set of residents $\hat{\mathcal{R}} = \bar{\mathcal{R}}'' \cup \{\mathcal{W}_\ell^k\}_{\ell \in [L], k \in [n^4]} \cup \mathcal{Y} \cup \mathcal{A} \cup \mathcal{U} \cup \bar{\mathcal{U}}$, and a set of hospitals $\hat{\mathcal{H}} = \bar{\mathcal{H}} \cup \mathcal{X} \cup \mathcal{V} \cup \bar{\mathcal{V}} \cup \mathcal{Z}$.

The sole purpose of the hospitals in $\mathcal{X}$ is to ensure that there are sufficient capacities for matching all the residents. The set of hospitals $\mathcal{Z}$ is introduced to make costly certain allocations of extra capacities. The sets $\mathcal{V}$ and $\bar{\mathcal{V}}$ are used to leverage stability and ensure that multiple allocations of extra seats yield sub-optimal solutions.

Capacity vector. Every hospital $j \in \hat{\mathcal{H}}$ has capacity one, i.e., $\hat{c}_j = 1$.

Preference lists. We now proceed to construct the preference lists in $\hat{\Gamma}$.

Given a resident $i_\ell \in \mathcal{R}'$ with $\ell \in [L]$, let $(j_{\sigma_1}, j_{\sigma_2}), j_{\sigma_3}, \dots$ be her ranking of the hospitals in the original instance Γ . We provide the preference lists of the residents and hospitals in village $\mathcal{B}_\ell^k$ with $\ell \in [L]$ and $k \in [n^4]$, namely

$$\begin{aligned} w_{\ell,1}^k &: v_{\ell,1}^k, \bar{\mathcal{V}}_\ell^k \setminus \{\bar{v}_{\ell,1}^k\}, \mathcal{V}_+^{k+1}[n^2], j_{\sigma(\ell,1)}^k, \mathcal{Z}, \dots, \mathcal{X} \\ w_{\ell,2}^k &: v_{\ell,2}^k, \bar{\mathcal{V}}_\ell^k \setminus \{\bar{v}_{\ell,2}^k\}, \mathcal{V}_+^{k+1}[n^2 + 1], j_{\sigma(\ell,2)}^k, \mathcal{Z}, \dots, \mathcal{X} \\ w_{\ell,e}^k &: v_{\ell,e}^k, \bar{\mathcal{V}}_\ell^k \setminus \{\bar{v}_{\ell,e}^k\}, \mathcal{V}_+^{k+1}[(n^2 - 1)e + 2], j_{\sigma(\ell,e)}^k, \mathcal{Z}, \dots, \mathcal{X} \\ &e \in \{3, \dots, n\} \\ y_\ell^k &: \bar{v}_{\ell,2}^k, \bar{v}_{\ell,1}^k, \bar{v}_{\ell,3}^k, \dots, \bar{v}_{\ell,n}^k, \mathcal{V}_+^{k+1}, \mathcal{Z}, \dots, \mathcal{X} \\ v_{\ell,e}^k &: u_{\ell,e}^k, \mathcal{A}, w_{\ell,e}^k, \{w_{\ell,e}^{k'}\}_{k' \neq k}, \{\mathcal{W}_\ell^k\}_{\ell,k} \setminus \{w_{\ell,e}^k\}_{k'}, \dots, \mathcal{Y} \\ &e \in [n] \\ u_{\ell,e}^k &: v_{\ell,e}^k, \dots & e \in [n] \\ \bar{v}_{\ell,e}^k &: \bar{u}_{\ell,e}^k, \mathcal{A}, \mathcal{W}_\ell^k \setminus \{w_{\ell,e}^k\}, y_\ell^k, \{\mathcal{W}_\ell^{k'}\}_{\ell,k' \neq k} \setminus \{w_{\ell,e}^{k'}\}_{k' \neq k}, \dots, \mathcal{Y} \setminus \{y_\ell^k\} \\ &e \in [n] \\ \bar{u}_{\ell,e}^k &: \bar{v}_{\ell,e}^k, \dots & e \in [n] \end{aligned}$$

where $j_{\sigma(\ell,e)}$ is the hospital listed in position e by resident i_ℓ , and $\mathcal{V}_+^k$ is the set of hospitals $\mathcal{V}$ listed as follows: $(\mathcal{V}^k, \dots, \mathcal{V}^{n^4}, \mathcal{V}^1, \dots, \mathcal{V}^{k-1})$, for $k \in [n^4]$; note that when $k \geq n^4$ then we take $\mathcal{V}_+^k$ for $k \pmod{n^4}$. Moreover, we denote by $\mathcal{V}_+^k[n]$ the first n elements of the ordered set $\mathcal{V}_+^k$. The purpose of positioning set $\mathcal{V}_+^{k+1}$ in the preference lists of the residents $w_{\ell,e}^k$, ($e \in [n]$), is to ensure that we can mimic $\text{rank}_{i_\ell}(j)$, $j \in \mathcal{H}$, of the original instance while also ensuring that multiple allocation of extra capacities to the same hospital are sub-optimal. The symbol “...” means that the agents on the other side of the bipartition not explicitly listed are ranked strictly and arbitrarily. Now, we present the preference list of a hospital $j_e^k \in \bar{\mathcal{H}}$, which is a

modification of the preference list of hospital $j_e \in \mathcal{H}$. We modify the original preference list of j_e by substituting every resident $i_\ell \in \mathcal{R}'$, $\ell \in [L]$, with resident $w_{\ell,r}^k$ where r is the position of hospital j_e in the preference list of resident i_ℓ (e.g., $r = 2$ if j_e is the second hospital listed in the tie at the head of the preference list of i_ℓ); at the head of the preference list of j_e^k , we position $\mathcal{A}$. Concerning the residents in $\mathcal{R}''$, we substitute every $i_\ell \in \mathcal{R}''$, ($\ell \in \{L+1, \dots, n\}$), with i_ℓ^k . Then, hospital j_h^k ranks arbitrarily a strict ordering of the remaining residents, and at the end $\mathcal{Y}$. We introduce the preference list of a resident $i_\ell^k \in \mathcal{R}''$ by modifying the strict preference list of the corresponding resident $i_\ell \in \mathcal{R}''$: $j_{\sigma(\ell,1)}, \dots, j_{\sigma(\ell,n)}$ in Γ . The preference list of i_ℓ^k in our new instance $\hat{\Gamma}$ is $\mathcal{V}_+^{k+1}[n^2 + 2], j_{\sigma_1}^k, \mathcal{V}_+^{k+1}[n^2 + 2, 1], j_{\sigma_2}^k, \mathcal{V}_+^{k+1}[n^2 + 2, 2], \dots, \mathcal{V}_+^{k+1}[n^2 + 2, n-1], j_{\sigma_n}^k, \mathcal{Z} \dots, \mathcal{X}$ where $\mathcal{V}_+^{k+1}[b, c]$ is the set of $n^2 - 1$ hospitals from $\mathcal{V}_+^{k+1}$ that is listed from index $b + (n^2 - 1) \cdot c$ onward. Note that if we go over the limit, we consider again the modulo class. The preference lists of residents in $\mathcal{A}$ is $a_p : z_p, \dots$; and the preference list of the hospitals in $\mathcal{Z}$ is $z_p : a_p, \dots$. The preference lists of hospitals in $\mathcal{X}$ are arbitrary. Finally, it is worth noting that every stable matching in $\hat{\Gamma}$ will include the following pairs: $\{(a_p, z_p)\}_{p \in [n^6]}$, and $\{(u_{\ell,e}^k, v_{\ell,e}^k), (\bar{u}_{\ell,e}^k, \bar{v}_{\ell,e}^k)\}_{\ell \in [L], e \in [n], k \in [n^4]}$, as they rank each other first.

Target value and budget. Consider an instance of MIN-w SMT with a given target K , where L is the number of residents with a tie at the head of their preference list. We define the target $\bar{K} = n^6 \cdot K + n^6 + 4n^5L + 3n^5 - n^4L$ and budget $B = L \cdot n^5$ for MIN-Avg_{EXP} HR. Note that $\hat{\mathcal{R}} = \mathcal{R}'' \cup \{w_{\ell,r}^k\}_{\ell \in [L], k \in [n^4]} \cup \mathcal{Y} \cup \mathcal{A} \cup \mathcal{U} \cup \bar{\mathcal{U}}$ and $\hat{\mathcal{H}} = \bar{\mathcal{H}} \cup \mathcal{X} \cup \mathcal{V} \cup \bar{\mathcal{V}} \cup \mathcal{Z}$. Therefore, the instance $\hat{\Gamma}$ consists of $((n-L)n^4 + L(n+1)n^4 + n^6 + 2n^5L) + (n^5 + n^7 + 2n^5L + n^6) = n^7 + 2n^6 + n^5(2+5L)$ residents and hospitals, which is $O(n^7)$; therefore the construction can be done in polynomial-time. Finally, if no extra seats are assigned in our new instance, the residents will be matched to $\mathcal{X}$, leading to a higher average hospital rank.

3.2. Useful Lemmata

Next, we prove that the reduction previously introduced maps a weakly stable matching of SMT into a stable matching of HR, where the average rank of the new matching is proportional to the average rank of the one in SMT.

Lemma 3.2. *Let M be a weakly stable matching in Γ with $\text{AvgRank}(M) = K_M$. There is an allocation $\mathbf{t}$ respecting the budget $B = L \cdot n^5$ and a stable matching $\hat{M}_t$ in $\hat{\Gamma}_t = \langle \hat{\mathcal{R}}, \hat{\mathcal{H}}, \hat{\succ}, \hat{\mathbf{c}} + \mathbf{t} \rangle$ with $\text{AvgRank}(\hat{M}_t) = n^6 \cdot K_M + n^6 + 4n^5L + 3n^5 - n^4L$.*

Proof. Let M be a (complete) weakly stable matching in Γ . Let $\text{index}_{i_\ell}(j)$ be the index of hospital j in resident i_ℓ preference list, i.e., $\text{index}_{i_\ell}(j)$ is equal to 1 or 2 if j is ranked first, otherwise is equal to the rank of j . Define the set of indices in M as $\text{Idx}(M) = \{(\ell, r) : r = \text{index}_{i_\ell}(j), (i_\ell, j) \in M \cap (\mathcal{R}' \times \mathcal{H})\}$. The set $\text{Idx}(M)$ contains the information of the pairs $\mathcal{R}' \times \mathcal{H}$ that are matched in M . Given M and Idx , we now define the following sets that will be helpful in this proof. For every $k \in [n^4]$, we define the set of residents $\mathcal{W}_M^k = \{w_{\ell,r}^k \in \mathcal{W} : (\ell, r) \in \text{Idx}(M)\}$, and define the sets of hospitals $\bar{\mathcal{V}}_M^k = \{\bar{v}_{\ell,r}^k \in \bar{\mathcal{V}} : (\ell, r) \in$

$\text{Idx}(M)\}$, and $\mathcal{V}_M^k = \{v_{\ell,e}^k \in \mathcal{V} : (\ell, r) \in \text{Idx}(M), e \in [n] \setminus \{r\}\}$. Now, we provide an allocation of extra seats $\mathbf{t}$ with a total budget $B = L \cdot n^5$ and a stable matching $\hat{M}_t$ in $\hat{\Gamma}_t$.

Allocation of extra seats. For every $k \in [n^4]$, we assign n extra seats to each hospital in $\bar{\mathcal{V}}_M^k \cup \mathcal{V}_M^k$. For the rest of the hospitals, we assign no extra capacities. Formally, for $k \in [n^4]$, we have $t_v = 1$ if $v \in \bar{\mathcal{V}}_M^k \cup \mathcal{V}_M^k$, and $t_v = 0$ otherwise. Since $L = |\mathcal{R}'|$, all of the extra seats $B = L \cdot n^5$ are used.

Matching. For each $(\ell, r) \in \text{Idx}(M)$ with j such that $r = \text{index}_{i_\ell}(j)$ in Γ , we match the following pairs in $\hat{M}_t$: $(w_{\ell,r}^k, j^k)$, $(y_\ell^k, \bar{v}_{\ell,r}^k)$, and $(w_{\ell,e}^k, v_{\ell,e}^k)$ for $e \in [n] \setminus \{r\}$ and $k \in [n^4]$. Note that if j is ranked first by i_ℓ , the hospital is listed first or second in the tie. If j is listed first, then $r = 1$ and we match the pairs $\{(w_{\ell,1}^k, j^k)\}_{k \in [n^4]}$, otherwise, $r = 2$ and we match the pairs $\{(w_{\ell,2}^k, j^k)\}_{k \in [n^4]}$. For each $(i, j) \in M$ with $i \in \mathcal{R}''$, we match the pairs $\{(i^k, j^k)\}_{k \in [n^4]}$ in $\hat{M}_t$, where j^k is a copy of j in $\bar{\mathcal{H}}$. Formally, $\hat{M}_t = \{(i^k, j^k) : (i, j) \in M \cap (\mathcal{R}'' \times \mathcal{H}) \text{ and } k \in [n^4]\} \cup \{(w_{\ell,r}^k, j^k) : r = \text{index}_{i_\ell}(j), (i_\ell, j) \in M \cap (\mathcal{R}' \times \mathcal{H}) \text{ and } k \in [n^4]\} \cup \{(y_\ell^k, \bar{v}_{\ell,r}^k) : (\ell, r) \in \text{Idx}(M) \text{ and } k \in [n^4]\} \cup \{(w_{\ell,e}^k, v_{\ell,e}^k) : (\ell, r) \in \text{Idx}(M), e \in [n] \setminus \{r\} \text{ and } k \in [n^4]\} \cup \{(u_{\ell,e}^k, v_{\ell,e}^k), (\bar{u}_{\ell,e}^k, \bar{v}_{\ell,e}^k)\}_{\ell \in [L], e \in [n], k \in [n^4]} \cup \{(a_p, z_p)\}_{p \in [n^6]}$.

The only hospitals that are matched to two residents are those that receive one extra capacity according to vector $\mathbf{t}$. All the other hospitals are matched at most to one resident. Therefore, $\hat{M}_t$ is a matching. Let us verify that $\hat{M}_t$ is a stable matching in $\hat{\Gamma}_t$. First, note that residents $i^k \in \mathcal{R}''$ and hospitals $j^k \in \bar{\mathcal{H}}$ cannot create blocking pairs because of their stability in M . Now, we check the stability of the pairs in each $\mathcal{B}_\ell^k$, where $i_\ell \in \mathcal{R}'$ with $\ell \in [L]$, and $k \in [n^4]$. The pairs matched in $\mathcal{B}_\ell^k$ are $(w_{\ell,r}^k, j^k)$, $(w_{\ell,e}^k, v_{\ell,e}^k)$, $(y_\ell^k, \bar{v}_{\ell,r}^k)$ and $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}_{p \in [n] \setminus \{r\}}$ where $e \in [n] \setminus \{r\}$.

The agents in the pair $(w_{\ell,r}^k, j^k)$ are not part of any blocking pair; in fact, $w_{\ell,r}^k$ cannot be matched to any of the hospitals in $\{v_{\ell,r}^k\} \cup \bar{\mathcal{V}}_\ell^k \setminus \{\bar{v}_{\ell,r}^k\}$ because they have capacity one and they are all matched to their most favorite resident in $\mathcal{U} \cup \bar{\mathcal{U}}$. Also j^k cannot be part of a blocking pair. Indeed, all the residents $w_{\ell',r'}^k$ ranked in the hospital's preference list before $w_{\ell,r}^k$ are matched to hospitals of the form $v_{\ell',r'}^k$ that they rank first or to another hospital j'^k they prefer over j^k (due to the stability of M in Γ). The overall structure of the village allows to mimic the original matching while adding a fixed cost of the village to the total cost of the matching. We display this idea in Figure 1. If i_ℓ and j are matched in the instance MIN-w SMT, then, in the reduced instance, we match the copy $w_{\ell,r}^k$ to the hospital-copy j^k ; note that the rest of the village $\mathcal{B}_\ell^k$ is matched to itself. Remarkably, we can match the remaining residents of the village $\mathcal{B}_\ell^k$ to the hospitals in the same village in a stable way.

For $e \in [n] \setminus \{r\}$, $w_{\ell,e}^k$ ranks $v_{\ell,e}^k$ first, and $v_{\ell,e}^k$ has capacity two and ranks $w_{\ell,e}^k$ second. Hence the $n-1$ pairs $(w_{\ell,e}^k, v_{\ell,e}^k)$ do not include any agent who may be part of a blocking pair.

If $r = 2$, then y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ first, and $\bar{v}_{\ell,r}^k$ cannot be matched to any of the residents in $\mathcal{W}_\ell^k \setminus \{w_{\ell,r}^k\}$ because of the previous point. Moreover, $\bar{v}_{\ell,r}^k$ has capacity two and is also matched to

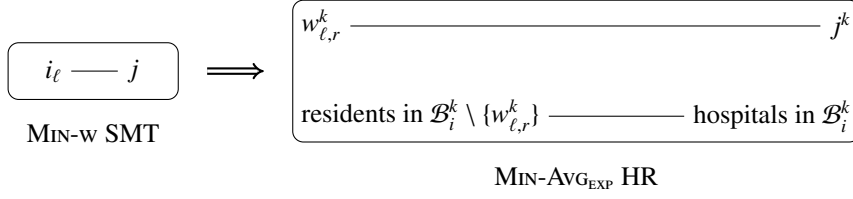


Figure 1: Edges denote matches: (i_ℓ, j) in MIN-w SMT maps to a stable pair $(w_{\ell,r}^k, j^k)$ in MIN-AvgExp HR via the village $\mathcal{B}_\ell^k$.

$\bar{u}_{\ell,r}^k$. Therefore, the agents in the pair $(y_\ell^k, \bar{v}_{\ell,r}^k)$ do not take part in a blocking pair when $r = 2$. If $r \neq 2$, then y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ second or more, and y_ℓ^k cannot be matched to $\bar{v}_{\ell,q}^k$ for $q \in \{2, 1, 3, \dots, r-1\}$: Each $\bar{v}_{\ell,q}^k$ has capacity one and it is already matched to $\bar{u}_{\ell,q}^k$. On the other hand, $\bar{v}_{\ell,r}^k$ cannot create a blocking pair with any of the residents in $\mathcal{W}_\ell^k \setminus \{w_{\ell,r}^k\}$ because these residents are matched to their most preferred hospitals. Then, the agents in the pair $(y_\ell^k, \bar{v}_{\ell,r}^k)$ do not create blocking pairs for $r \neq 2$.

The pairs $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}_{p \in [n]}$ do not involve any agent who could be part of a blocking pair. First, each of these listed pairs matches two agents that rank each other first. Second, for those hospitals with capacity two, from the previous points, we conclude that there are no residents with which they could create a blocking pair.

Therefore, $\hat{M}_t$ is a stable matching in $\hat{\Gamma}_t$. Next, we compute the average hospital rank in M and $\hat{M}_t$. In M , we can distinguish whether a resident is in $\mathcal{R}''$ or $\mathcal{R}'$ and we can distinguish if a resident is matched to a hospital ranked first or not. Let G'' be the average hospital rank of residents in $\mathcal{R}''$, K'_t be the average hospital rank of the residents in $\mathcal{R}'$ that are matched to a hospital in their ties, and K'_s be the average hospital rank of the residents in $\mathcal{R}'$ that are matched to a hospital they rank third or more. Note that K'_t is also the number of residents from $\mathcal{R}'$ matched to a hospital they rank first. The average hospital rank of M is $K_M = G'' + K'_t + K'_s$.

We now prove that, given $k \in [n^4]$ and $(\ell, r) \in \text{Idx}(M)$, the average rank of the residents in village $\mathcal{B}_\ell^k$ is: $(n + n^2 \cdot r + 3) + (n - 1) + (2n)$. The first term comes from the residents matched in the two pairs $(w_{\ell,r}^k, j^k)$ and $(y_\ell^k, \bar{v}_{\ell,r}^k)$: if $r = 1$, then $w_{\ell,1}^k$ ranks j^k at position $n + n^2 + 1$ and y_ℓ^k ranks $\bar{v}_{\ell,1}^k$ at position 2, thus summing to $n + n^2 + 3$; if $r = 2$ then $w_{\ell,2}^k$ ranks j^k at position $n + n^2 + 2$ and y_ℓ^k ranks $\bar{v}_{\ell,2}^k$ at position 1, thus summing to $n + n^2 + 3$; if $r > 2$, then $w_{\ell,r}^k$ ranks j^k at position $n + (n^2 - 1)r + 2 + 1$ and y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ at position r , thus summing to $n + n^2 + 3$. The second term comes from the $(n - 1)$ residents matched in the pairs $\{(w_{\ell,e}^k, v_{\ell,e}^k)\}_{e \in [n] \setminus \{r\}}$. The third term comes from the $2n$ residents matched in the pairs $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}$. Note that $(n + n^2 \cdot r + 3) + (n - 1) + (2n) = n^2 \cdot r + 3 + (4n - 1)$. On the other hand, each resident i_ℓ^k for $\ell \in \{L + 1, \dots, n\}$ is matched to a hospital j^k , and w.l.o.g. j is ranked r by i_ℓ in Γ ; therefore, i_ℓ^k ranks j^k at position $(n^2 + 2) + 1 + (r - 1) \cdot (n^2 - 1) + (r - 1) = n^2 \cdot r + 3$, where the first term comes from the cardinality of the first set ranked by i_ℓ^k , the second term comes from the first j ranked by i_ℓ , the third term comes from the number of sets of the form $\mathcal{V}_+^{k+1}[n^2 + 2, e]$ (each of cardinality $n^2 - 1$) positioned before j^k ,

and the last term is the number of hospitals j^k ranked before j^k .

Therefore, the residents in copy k have an average rank of $[n^2 \cdot G'' + 3(n - L)] + n^2 \cdot (K'_t + K'_s) + [3 + (4n - 1)] \cdot (L) = n^2 \cdot K_M + 3(n - L + L) + (4n - 1)L = n^2 \cdot K_M + 4nL + 3n - L$, where the first term comes from the residents i_ℓ^k for $\ell \in \{L + 1, \dots, n\}$ (note that when we sum the contribution of each such resident, we have the sum $\sum_{\ell \in \{L+1, \dots, n\}} n^2 \cdot r_\ell + 3$, where r_ℓ is $\text{index}_{i_\ell}(M(i_\ell))$), the second term comes from the residents $\{w_{\ell,r}^k, y_\ell^k\}$ for $(\ell, r) \in \text{Idx}(M)$, and the third term comes from the remaining residents in the villages indexed with k and the constant 3 multiplied by the number of students in villages matched to a hospital j^k . Note that when we sum the contribution of each resident in every village that is not matched to a hospital $v_{\ell,e}^k$, we have the sum $\sum_{\ell \in \{L+1, \dots, n\}} n^2 \cdot r_\ell + 3$, where r_ℓ is $\text{index}_{i_\ell}(M(i_\ell))$. Since $k \in [n^4]$, then $\text{AvgRank}(\hat{M}_t)$ is $n^4 \cdot (n^2 \cdot K_M + 4nL + 3n - L) + n^6$, where the last term comes from the residents in $\mathcal{A}$. Hence $\text{AvgRank}(\hat{M}_t) = n^6 \cdot K_M + n^6 + 4n^5L + 3n^5 - n^4L$. $\square$

In Appendix Appendix A.3, we provide an example of an instance showcasing the details of a village.

Now, we show that the allocation vector and the stable matching from Lemma 3.2 correspond to the solution with the minimum average hospital rank, as long as the original matching of Γ is of minimum average hospital rank.

Lemma 3.3. *Consider a weakly stable matching M in Γ of minimum average hospital rank. Then, the allocation $\mathbf{t}$ and the stable matching $\hat{M}_t$ from Lemma 3.2 are a solution of minimum average hospital rank for $\hat{\Gamma}$ when $B = L \cdot n^5$.*

The proof can be found in Appendix Appendix A.2.

3.3. MIN-AvgExp HR is NP-complete

Proof of Theorem 3.1. MIN-AvgExp HR is clearly in NP since given $\mathbf{t}$ and a matching $\hat{M}_t$ in instance $\hat{\Gamma}_t$, we can verify in polynomial-time whether $\hat{M}_t$ is stable, whether the budget B is allocated and whether its objective value is less than the target value. We now show that MIN-AvgExp HR is NP-complete.

From Lemma 3.1, we know that MIN-w SMT is NP-complete. Consider the reduction given in Section 3.1. In the constructed instance $\hat{\Gamma}$ of MIN-AvgExp HR, we set the budget to $B = L \cdot n^5$ and the target value to $\bar{K} = n^6 \cdot K + n^6 + 4n^5L + 3n^5 - n^4L$, where K is the target value of the instance of MIN-w SMT.

First, suppose that the answer to the instance of MIN-w SMT is NO, i.e., there is no weakly stable matching M with an average hospital rank less or equal than K . Let M be a weakly stable matching with minimum average hospital rank K_M ; note that $K_M > K$. Next, we prove that there is no allocation of

extra seats and a stable matching in the respective instance $\hat{\Gamma}$ of MIN-AVG_{EXP} HR with an objective value less or equal than $n^6 \cdot K + n^6 + 4n^5L + 3n^5 - n^4L$. Indeed, in Lemma 3.2, we show that there is an allocation $\mathbf{t}$ and a stable matching $\hat{M}_{\mathbf{t}}$ with $\text{AvgRank}(\hat{M}_{\mathbf{t}}) = n^6 \cdot K_M + n^6 + 4n^5L + 3n^5 - n^4L$. In Lemma 3.3, we prove that this is a minimum average hospital rank for $\hat{\Gamma}$ since M is an optimal matching. Therefore, $n^6 \cdot K_M + n^6 + 4n^5L + 3n^5 - n^4L > n^6 \cdot K + n^6 + 4n^5L + 3n^5 - n^4L = \bar{K}$, which means that the answer for the instance of MIN-AVG_{EXP} HR is also NO.

On the other hand, consider a YES instance of MIN-w SMT. Then, there is a weakly stable matching M with an average hospital rank of $K_M \leq K$. Therefore, the allocation $\mathbf{t}$ and the stable matching $\hat{M}_{\mathbf{t}}$ in $\hat{\Gamma}$ constructed in Lemma 3.2 have an objective value of $K_M + n^6 + 4n^5L + 3n^5 - n^4L \leq n^6 \cdot K + n^6 + 4n^5L + 3n^5 - n^4L = \bar{K}$. Hence, the instance of MIN-AVG_{EXP} HR has a YES answer.

Let us prove now that, for any ε , MIN-AVG_{EXP} HR OPT is not approximable within a factor of $(\bar{n})^{1/6-\varepsilon}$, where $\bar{n}$ is the number of hospitals in the reduced instance, unless P=NP. Consider an instance Γ of MIN-w SMT with n hospitals and n residents, where $L \geq 1$ of the residents have a tie in their preference list. Let M^{yes} and M^{no} be the stable matchings of minimum average hospital rank for the cases in which the answer of the decision problem MIN-w SMT is YES and NO, respectively. Lemma 3.1 implies that, for any $\varepsilon > 0$, $\text{AvgRank}(M^{no}) \geq n^{1-\varepsilon} \cdot \text{AvgRank}(M^{yes})$. Now, consider the reduction presented in Section 3.1 from instance Γ to an instance $\hat{\Gamma}$ of MIN-AVG_{EXP} HR. Lemma 3.3 implies that there are allocations $\mathbf{t}$ and $\mathbf{t}'$, and matchings $\hat{M}_{\mathbf{t}}^{yes}$ and $\hat{M}_{\mathbf{t}'}^{no}$ for the respective YES and NO answers of MIN-AVG_{EXP} HR such that $\text{AvgRank}(\hat{M}_{\mathbf{t}}^{yes}) = n^6 \cdot \text{AvgRank}(M^{yes}) + n^6 + 4n^5L + 3n^5 - n^4L$ and $\text{AvgRank}(\hat{M}_{\mathbf{t}'}^{no}) = n^6 \cdot \text{AvgRank}(M^{no}) + n^6 + 4n^5L + 3n^5 - n^4L$. Then, for any $\varepsilon > 0$, we have $\frac{\text{AvgRank}(\hat{M}_{\mathbf{t}'}^{no})}{\text{AvgRank}(\hat{M}_{\mathbf{t}}^{yes})} = \frac{n^6 \cdot \text{AvgRank}(M^{no}) + n^6 + 4n^5L + 3n^5 - n^4L}{n^6 \cdot \text{AvgRank}(M^{yes}) + n^6 + 4n^5L + 3n^5 - n^4L} \geq c \cdot \frac{\text{AvgRank}(M^{no})}{\text{AvgRank}(M^{yes})} \geq c \cdot n^{1-\varepsilon} \geq c \cdot \left(\frac{\bar{n}}{2}\right)^{\frac{1-\varepsilon}{6}}$ for some constant $c > 0$. This completes the proof. $\square$

Acknowledgments

This work was funded by IVADO and Fonds de recherche du Québec (FRQ) with the grant number “FRQ-IVADO Research Chair in Data Science for Combinatorial Game Theory”, and Natural Sciences and Engineering Research Council of Canada (NSERC) with the grant number “discovery grant 2019-04557”. Part of the work was conducted when the third author was Canada Excellence Research Chair (CERC) “Canada Excellence Research Chair in Data Science for Real-Time Decision-Making” at Polytechnique Montréal, the generous support of the CERC grant being warmly acknowledged.

References

[1] Abe, K., Komiyama, J., Iwasaki, A., 2022. Anytime capacity expansion in medical residency match by Monte Carlo tree search. arXiv preprint arXiv:2202.06570.

[2] Afacan, M.O., Dur, U., Van der Linden, M., 2024. Capacity design in school choice. *Games and Economic Behavior* 146, 277–291.

[3] Aziz, H., Brandl, F., 2021. Efficient, fair, and incentive-compatible healthcare rationing, in: *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 103–104.

[4] Biró, P., 2008. Student admissions in Hungary as Gale and Shapley envisaged. University of Glasgow Technical Report TR-2008-291.

[5] Bobbio, F., Carvalho, M., Lodi, A., Rios, I., Torrico, A., 2021. Capacity planning in stable matching: An application to school choice. arXiv preprint arXiv:2110.00734.

[6] Bobbio, F., Carvalho, M., Lodi, A., Rios, I., Torrico, A., 2023. Capacity planning in stable matching: An application to school choice, in: *Proceedings of the 22nd ACM Conference on Economics and Computation*.

[7] Bobbio, F., Carvalho, M., Lodi, A., Torrico, A., 2022. Capacity variation in the many-to-one stable matching. arXiv preprint arXiv:2205.01302.

[8] Chen, J., Csáji, G., 2023. Optimal capacity modification for many-to-one matching problems, in: *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, pp. 2880–2882.

[9] Delacretaz, D., Kominers, S.D., Teytelboym, A., 2016. Refugee resettlement.

[10] Gale, D., Shapley, L.S., 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 9–15.

[11] Gale, D., Sotomayor, M., 1985. Some remarks on the stable matching problem. *Discrete Applied Mathematics* 11, 223–232.

[12] Kelso Jr, A.S., Crawford, V.P., 1982. Job matching, coalition formation, and gross substitutes. *Econometrica*, 1483–1504.

[13] Kominers, S.D., 2019. Respect for improvements and comparative statics in matching markets. Technical Report. Working Paper.

[14] Kumano, T., Kurino, M., et al., 2022. Quota Adjustment Process. Technical Report. Institute for Economics Studies, Keio University.

[15] Manlove, D.F., Irving, R.W., Iwama, K., Miyazaki, S., Morita, Y., 2002. Hard variants of stable marriage. *Theoretical Computer Science* 276, 261–279.

[16] Nguyen, T., Vohra, R., 2018. Near-feasible stable matchings with couples. *American Economic Review* 108, 3154–69.

[17] Ranjan, K., Nasre, M., Nimbhorkar, P., 2024. Optimal capacity modification for strongly stable matchings. arXiv preprint arXiv:2411.10284.

[18] Romm, A., 2014. Implications of capacity reduction and entry in many-to-one stable matching. *Social Choice and Welfare* 43, 851–875.

[19] Roth, A.E., 1984. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of political Economy* 92, 991–1016.

[20] Sönmez, T., 1997. Manipulation via capacities in two-sided matching markets. *Journal of Economic theory* 77, 197–204.

[21] Yahiro, K., Yokoo, M., 2020. Game theoretic analysis for two-sided matching with resource allocation, in: *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1548–1556.

Appendix A. Missing proofs

The following problem is useful for the missing proof that we provide in this Appendix.

Problem 3 (MAX-CARD HRTI).

INSTANCE: An HRTI instance $\Gamma = \langle \mathcal{R}, \mathcal{H}, >, \mathbf{c} \rangle$ with $c_j = 1$ for all $j \in \mathcal{H}$, $|\mathcal{H}| = |\mathcal{R}|$ and a non-negative integer target value $K \in \mathbb{Z}_+$.

QUESTION: Is there a weakly stable matching M such that $|M| \geq K$?

Recall that HRTI corresponds to the problem with ties and incomplete preference lists. Manlove et al. [15] proved that MAX-CARD HRTI is NP-complete. As the next remark states, this result holds even if ties are at the head of the preference

list, only on one side of it, at most one tie per list, and each tie is of length 2.

Remark 1. After the proof of Lemma 1 in [15], the authors showed that the problem MAX-CARD HRTI can be simplified to the case in which ties are only on one side of the bipartition and are at the end of the preference list. Since the ties of the new instance created in Lemma 1 from [15] are of length at most two, we can use the same reasoning to assume instead, without loss of generality, that in an instance of MAX-CARD HRTI and the corresponding MIN-w SMT instance of Lemma 3.1 ties occur only at the head of a preference list.

Appendix A.1. Proof of Lemma 3.1

In this section, we prove that MIN-w SMT is NP-complete and its optimization version cannot be approximated within a constant factor. The proof is inspired by the proof of Theorem 7 in [15]. The result in [15] is stated in the traditional notation of the stable marriage problem where both sides are defined as women and men, instead of residents and hospitals, respectively. To keep coherence with the previous work, for this proof we also denote both sides as women and men.

Proof of Lemma 3.1. Clearly, MIN-w SMT is in NP. Given $\varepsilon > 0$, let $a = \lceil (3/\varepsilon) \rceil$. From Theorem 2 in [15], we know that, when ties occur on the women's side only, and each tie has length two, MAX-CARD HRTI is NP-complete. Consider an instance of Problem 3 with $\mathcal{H} = \{m_1, m_2, \dots, m_n\}$ and $\mathcal{R} = \{w_1, w_2, \dots, w_n\}$. We assume that the target value K is equal to n , since it was shown that even for this target value the problem is NP-complete. Let O_h (resp. R_h) denote the preference list of man m_h (resp. woman w_h) for $h \in [n]$. Next, we build an instance of MIN-w SMT. Let $C := n^{a-1}$, then

- the set of men is $\mathcal{H}' = \mathcal{H}^0 \cup \left(\bigcup_{h=1}^C \mathcal{H}^h \right)$ with $\mathcal{H}^0 = \{m_1^0, m_2^0, \dots, m_n^0\}$ and $\mathcal{H}^h = \{m_1^h, m_2^h, \dots, m_n^h\}$ for $h \in [C]$;
- the set of women is $\mathcal{R}' = \mathcal{R}^0 \cup \left(\bigcup_{h \in [C]} \mathcal{R}^h \right)$ with $\mathcal{R}^0 = \{w_1^0, w_2^0, \dots, w_n^0\}$ and $\mathcal{R}^h = \{w_1^h, w_2^h, \dots, w_n^h\}$ for $h \in [C]$;
- for each $h \in [n]$ and $s \in [C]$, let O_h^s be the preference list obtained from O_h by replacing woman w_k in O_h by the corresponding woman w_k^s , for every $k \in [n]$. We refer to the women in O_h^s as the proper women for m_h^s . Similarly, we define R_h^s and the proper men for w_h^s . The preference lists for $\mathcal{H}'$ and $\mathcal{R}'$ are

$$\begin{array}{ll} m_h^0 : w_h^0 \dots & h \in [n^a] \\ m_h^s : O_h^s, \mathcal{R}^0 \dots & h \in [n], s \in [C] \\ w_h^0 : m_h^0 \dots & h \in [n^a] \\ w_h^s : R_h^s, \mathcal{H}^0 \dots & h \in [n], s \in [C] \end{array}$$

where the dots “...” in the preference lists mean that the remaining agents on the other side of the bipartition are ranked strictly and arbitrarily, and the sets mean that the agents within are ranked according to their indices;

- the target value is $K' = \lfloor (n^{a+2})/2 \rfloor$.

Our MIN-w SMT instance comprises $2n^a$ men and $2n^a$ women, so that $\bar{n} := 2n^a$. Note also that the only ties in MIN-w SMT occur in the preference lists of women w_h^s for $h \in [n]$, $s \in [C]$. Moreover, there is at most one tie per list, and each tie has length 2.

Suppose that we have a YES instance for MAX-CARD HRTI, i.e., there is a stable matching M with $|M| = n$. We create a matching M' in MIN-w SMT as follows: For every $h \in [n^a]$, we add the pair (m_h^0, w_h^0) to M' , and for each $s \in [n]$, we add the pair (m_s^ℓ, w_k^ℓ) to M' for all $\ell \in [C]$, where $(m_s, w_k) \in M$. Note that M' is stable for our MIN-w SMT instance. We also have that

$$\text{AvgRank}(M') \leq n^a + n^{a-1}n^2 \leq \left\lfloor \frac{n^{a+2}}{2} \right\rfloor = K',$$

since, without loss of generality, we may choose $n \geq 3$. Therefore, the objective value in MIN-w SMT satisfies the target of K' .

On the other side, let us suppose that we have a NO instance for MAX-CARD HRTI, i.e., it does not have a stable matching of cardinality n . Then, in any stable matching M' of MIN-w SMT, it holds that, for every $s \in [C]$, there is some $h \in [n]$ for which w_h^s is not matched to one of her proper men. Nonetheless, in M' , m_h^0 and w_h^0 must be partners, for every $h \in [n^a]$. Therefore, there is some $h \in [n]$ such that $\text{rank}_{w_h^s}(M'(w_h^s)) > n^a$. Hence, $\text{AvgRank}(M') > n^{2a-1} > K'$ for any stable matching of our MIN-w SMT instance.

Therefore, the existence of a polynomial-time approximation algorithm for MIN-w SMT OPT whose approximation ratio is as good as $(2n^{2a-1})/n^{a+2} = 2n^{a-3}$ would give a polynomial-time algorithm for determining whether MAX-CARD HRTI has a stable matching in which everybody is matched (i.e., $K = n$). To conclude, we note that $2n^{a-3} = (2/2^{1-3/a})\bar{n}^{1-3/a} > \bar{n}^{1-3/a} > \bar{n}^{1-\varepsilon}$, which ends the proof. $\square$

Appendix A.2. Proof of Lemma 3.3

Proof of Lemma 3.3. Let M be a (complete) weakly stable matching in Γ . Let $\text{index}_{i_\ell}(j)$ be the index of hospital j in resident i_ℓ preference list, i.e., $\text{index}_{i_\ell}(j)$ is equal to 1 or 2 if j is ranked first, otherwise is equal to the rank of j . Define the set of indices in M as $\text{Idx}(M) = \{(\ell, r) : r = \text{index}_{i_\ell}(j), (i_\ell, j) \in M \cap (\mathcal{R}' \times \mathcal{H})\}$. The set $\text{Idx}(M)$ contains the information of the pairs $\mathcal{R}' \times \mathcal{H}$ that are matched in M . Given M and Idx , we now define the following sets that will be helpful in this proof. For every $k \in [n^4]$, we define the set of residents $\mathcal{W}_M^k = \{w_{\ell,r}^k \in \mathcal{W} : (\ell, r) \in \text{Idx}(M)\}$, and define the sets of hospitals $\overline{\mathcal{V}}_M^k = \{\overline{v}_{\ell,r}^k \in \overline{\mathcal{V}} : (\ell, r) \in \text{Idx}(M)\}$, and $\mathcal{V}_M^k = \{v_{\ell,e}^k \in \mathcal{V} : (\ell, r) \in \text{Idx}(M), e \in [n] \setminus \{r\}\}$. Now, we provide an allocation of extra seats $\mathbf{t}$ with a total budget $B = L \cdot n^5$ and a stable matching $\hat{M}_{\mathbf{t}}$ in $\hat{\Gamma}_{\mathbf{t}}$.

Allocation of extra seats. For every $k \in [n^4]$, we assign n extra seats to each hospital in $\overline{\mathcal{V}}_M^k \cup \mathcal{V}_M^k$. For the rest of the hospitals, we assign no extra capacities. Formally, for $k \in [n^4]$,

we have $t_v = 1$ if $v \in \overline{\mathcal{V}}_M^k \cup \mathcal{V}_M^k$, and $t_v = 0$ otherwise. Since $L = |\mathcal{R}'|$, all of the extra seats $B = L \cdot n^5$ are used.

Matching. For each $(\ell, r) \in \text{Idx}(M)$ with j such that $r = \text{index}_{i_\ell}(j)$ in Γ , we match the following pairs in $\hat{M}_t$: $(w_{\ell,r}^k, j^k)$, $(y_\ell^k, \bar{v}_{\ell,r}^k)$, and $(w_{\ell,e}^k, v_{\ell,e}^k)$ for $e \in [n] \setminus \{r\}$ and $k \in [n^4]$. Note that if j is ranked first by i_ℓ , the hospital is listed first or second in the tie. If j is listed first, then $r = 1$ and we match the pairs $\{(w_{\ell,1}^k, j^k)\}_{k \in [n^4]}$, otherwise, $r = 2$ and we match the pairs $\{(w_{\ell,2}^k, j^k)\}_{k \in [n^4]}$. For each $(i, j) \in M$ with $i \in \mathcal{R}''$, we match the pairs $\{(i^k, j^k)\}_{k \in [n^4]}$ in $\hat{M}_t$, where j^k is a copy of j in $\overline{\mathcal{H}}$. Formally, $\hat{M}_t = \{(i^k, j^k) : (i, j) \in M \cap (\mathcal{R}'' \times \mathcal{H}) \text{ and } k \in [n^4]\} \cup \{(w_{\ell,r}^k, j^k) : r = \text{index}_{i_\ell}(j), (i_\ell, j) \in M \cap (\mathcal{R}' \times \mathcal{H}) \text{ and } k \in [n^4]\} \cup \{(y_\ell^k, \bar{v}_{\ell,r}^k) : (\ell, r) \in \text{Idx}(M) \text{ and } k \in [n^4]\} \cup \{(w_{\ell,e}^k, v_{\ell,e}^k) : (\ell, r) \in \text{Idx}(M), e \in [n] \setminus \{r\} \text{ and } k \in [n^4]\} \cup \{(u_{\ell,e}^k, v_{\ell,e}^k), (\bar{u}_{\ell,e}^k, \bar{v}_{\ell,e}^k)\}_{\ell \in [L], e \in [n], k \in [n^4]} \cup \{(a_p, z_p)\}_{p \in [n^6]}$.

The only hospitals that are matched to two residents are those that receive one extra capacity according to vector $\mathbf{t}$. All the other hospitals are matched at most to one resident. Therefore, $\hat{M}_t$ is a matching. Let us verify that $\hat{M}_t$ is a stable matching in $\hat{\Gamma}_t$. First, note that residents $i^k \in \overline{\mathcal{R}}''$ and hospitals $j^k \in \overline{\mathcal{H}}$ cannot create blocking pairs because of their stability in M . Now, we check the stability of the pairs in each $\mathcal{B}_\ell^k$, where $i_\ell \in \mathcal{R}'$ with $\ell \in [L]$, and $k \in [n^4]$. The pairs matched in $\mathcal{B}_\ell^k$ are $(w_{\ell,r}^k, j^k)$, $(w_{\ell,e}^k, v_{\ell,e}^k)$, $(y_\ell^k, \bar{v}_{\ell,r}^k)$ and $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}_{p \in [n] \setminus \{r\}}$.

The agents in the pair $(w_{\ell,r}^k, j^k)$ are not part of any blocking pair; in fact, $w_{\ell,r}^k$ cannot be matched to any of the hospitals in $\{v_{\ell,r}^k\} \cup \overline{\mathcal{V}}_\ell^k \setminus \{\bar{v}_{\ell,r}^k\}$ because they have capacity one and they are all matched to their most favorite resident in $\mathcal{U} \cup \overline{\mathcal{U}}$. Also j^k cannot be part of a blocking pair. Indeed, all the residents $w_{\ell',r'}^k$ ranked in the hospital' preference list before $w_{\ell,r}^k$ are matched to hospitals of the form $v_{\ell',r'}^k$ that they rank first or to another hospital j^k they prefer over j^k (due to the stability of M in Γ).

For $e \in [n] \setminus \{r\}$, $w_{\ell,e}^k$ ranks $v_{\ell,e}^k$ first, and $v_{\ell,e}^k$ has capacity two and ranks $w_{\ell,e}^k$ second. Hence the $n - 1$ pairs $(w_{\ell,e}^k, v_{\ell,e}^k)$ do not include any agent who may be part of a blocking pair.

If $r = 2$, then y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ first, and $\bar{v}_{\ell,r}^k$ cannot be matched to any of the residents in $\mathcal{W}_\ell^k \setminus \{w_{\ell,r}^k\}$ because of the previous point. Moreover, $\bar{v}_{\ell,r}^k$ has capacity two and is also matched to $\bar{u}_{\ell,r}^k$. Therefore, the agents in the pair $(y_\ell^k, \bar{v}_{\ell,r}^k)$ do not take part in a blocking pair when $r = 2$. If $r \neq 2$, then y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ second or more, and y_ℓ^k cannot be matched to $\bar{v}_{\ell,q}^k$ for $q \in \{2, 1, 3, \dots, r-1\}$: Each $\bar{v}_{\ell,q}^k$ has capacity one and it is already matched to $\bar{u}_{\ell,q}^k$. On the other hand, $\bar{v}_{\ell,r}^k$ cannot create a blocking pair with any of the residents in $\mathcal{W}_\ell^k \setminus \{w_{\ell,r}^k\}$ because these residents are matched to their most preferred hospitals. Then, the agents in the pair $(y_\ell^k, \bar{v}_{\ell,r}^k)$ do not create blocking pairs for $r \neq 2$.

The pairs $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}_{p \in [n]}$ do not involve any agent who could be part of a blocking pair. First, each of these listed pairs matches two agents that rank each other first. Second, for those hospitals with capacity two, from the previous points, we conclude that there are no residents with which they could

create a blocking pair.

Therefore, $\hat{M}_t$ is a stable matching in $\hat{\Gamma}_t$. Next, we compute the average hospital rank in M and $\hat{M}_t$. In M , we can distinguish whether a resident is in $\mathcal{R}''$ or $\mathcal{R}'$ and we can distinguish if a resident is matched to a hospital ranked first or not. Let G'' be the average hospital rank of residents in $\mathcal{R}''$, K'_t be the average hospital rank of the residents in $\mathcal{R}'$ that are matched to a hospital in their ties, and K'_s be the average hospital rank of the residents in $\mathcal{R}'$ that are matched to a hospital they rank third or more. Note that K'_t is also the number of residents from $\mathcal{R}'$ matched to a hospital they rank first. The average hospital rank of M is $K_M = G'' + K'_t + K'_s$.

We now prove that, given $k \in [n^4]$ and $(\ell, r) \in \text{Idx}(M)$, the average rank of the residents in village $\mathcal{B}_\ell^k$ is: $(n + n^2 \cdot r + 3) + (n - 1) + (2n)$. The first term comes from the residents matched in the two pairs $(w_{\ell,r}^k, j^k)$ and $(y_\ell^k, \bar{v}_{\ell,r}^k)$: if $r = 1$, then $w_{\ell,1}^k$ ranks j^k at position $n + n^2 + 1$ and y_ℓ^k ranks $\bar{v}_{\ell,1}^k$ at position 2, thus summing to $n + n^2 + 3$; if $r = 2$ then $w_{\ell,2}^k$ ranks j^k at position $n + n^2 + 2$ and y_ℓ^k ranks $\bar{v}_{\ell,2}^k$ at position 1, thus summing to $n + n^2 + 3$; if $r > 2$, then $w_{\ell,r}^k$ ranks j^k at position $n + (n^2 - 1)r + 2 + 1$ and y_ℓ^k ranks $\bar{v}_{\ell,r}^k$ at position r , thus summing to $n + n^2 + 3$. The second term comes from the $(n - 1)$ residents matched in the pairs $\{(w_{\ell,e}^k, v_{\ell,e}^k)\}_{e \in [n] \setminus \{r\}}$. The third term comes from the $2n$ residents matched in the pairs $\{(u_{\ell,p}^k, v_{\ell,p}^k), (\bar{u}_{\ell,p}^k, \bar{v}_{\ell,p}^k)\}$. Note that $(n + n^2 \cdot r + 3) + (n - 1) + (2n) = n^2 \cdot r + 3 + (4n - 1)$. On the other hand, each resident i_ℓ^k for $\ell \in \{L + 1, \dots, n\}$ is matched to a hospital j^k , and w.l.o.g. j is ranked r by i_ℓ in Γ ; therefore, i_ℓ^k ranks j^k at position $(n^2 + 2) + 1 + (r - 1) \cdot (n^2 - 1) + (r - 1) = n^2 \cdot r + 3$, where the first term comes from the cardinality of the first set ranked by i_ℓ^k , the second term comes from the first j ranked by i_ℓ , the third term comes from the number of sets of the form $\mathcal{V}_\ell^{k+1}[n^2 + 2, e]$ (each of cardinality $n^2 - 1$) positioned before j^k , and the last term is the number of hospitals j^k ranked before j^k .

Therefore, the residents in copy k have an average rank of $[n^2 \cdot G'' + 3(n - L)] + n^2 \cdot (K'_t + K'_s) + [3 + (4n - 1)] \cdot (L) = n^2 \cdot K_M + 3(n - L + L) + (4n - 1)L = n^2 \cdot K_M + 4nL + 3n - L$, where the first term comes from the residents i_ℓ^k for $\ell \in \{L + 1, \dots, n\}$ (note that when we sum the contribution of each such resident, we have the sum $\sum_{\ell \in \{L + 1, \dots, n\}} n^2 \cdot r_\ell + 3$, where r_ℓ is $\text{index}_{i_\ell}(M(i_\ell))$), the second term comes from the residents $\{w_{\ell,r}^k, y_\ell^k\}$ for $(\ell, r) \in \text{Idx}(M)$, and the third term comes from the remaining residents in the villages indexed with k and the constant 3 multiplied by the number of students in villages matched to a hospital j^k . Note that when we sum the contribution of each resident in every village that is not matched to a hospital $v_{\ell,e}^k$, we have the sum $\sum_{\ell \in \{L + 1, \dots, n\}} n^2 \cdot r_\ell + 3$, where r_ℓ is $\text{index}_{i_\ell}(M(i_\ell))$. Since $k \in [n^4]$, then $\text{AvgRank}(\hat{M}_t)$ is $n^4 \cdot (n^2 \cdot K_M + 4nL + 3n - L) + n^6$, where the last term comes from the residents in $\mathcal{A}$. Hence $\text{AvgRank}(\hat{M}_t) = n^6 \cdot K_M + n^6 + 4n^5L + 3n^5 - n^4L$. $\square$

Appendix A.3. Village Example

We present an example of an instance of MIN-w SMT and provide its reduced instance to an instance of MIN-Avg_{EXP} HR. Consider an instance I with $K = 3$, $\mathcal{H} = \{j_1, j_2, j_3\}$ and $\mathcal{R} = \{i_1, i_2, i_3\}$. Every agent has capacity 1, and the preference lists are as follows.

$$\begin{array}{ll}
i_1 : (j_1, j_2), j_3 & j_1 : i_1, i_2, i_3 \\
i_2 : (j_3, j_2), j_1 & j_2 : i_2, i_1, i_3 \\
i_3 : j_2, j_1, j_3 & j_3 : i_1, i_3, i_2
\end{array}$$

Let us now provide the reduced instance of MIN-AVG_{EXP} HR. The residents are $\bar{\mathcal{R}}'' = \{i_3\}$, $\{\mathcal{W}_\ell^k\}_{\ell \in [L], k \in [n^4]} = \{w_{\ell,1}^k, w_{\ell,2}^k, w_{\ell,3}^k\}_{\ell \in \{1,2\}, k \in [81]}$, $\mathcal{Y} = \{y_\ell^k\}_{\ell \in \{1,2\}, k \in [81]}$, $\mathcal{A} = \{a_1, \dots, a_{729}\}$, $\mathcal{U} = \{u_{\ell,1}^k, u_{\ell,2}^k, u_{\ell,3}^k\}_{\ell \in \{1,2\}, k \in [81]}$, $\bar{\mathcal{U}} = \{\bar{u}_{\ell,1}^k, \bar{u}_{\ell,2}^k, \bar{u}_{\ell,3}^k\}_{\ell \in \{1,2\}, k \in [81]}$. The hospitals are $\bar{\mathcal{H}} = \{j_1^k, j_2^k, j_3^k\}_{k \in [81]}$, $\mathcal{X} = \{x_1, \dots, x_{2187}\}$, $\mathcal{V} = \{v_{\ell,1}^k, v_{\ell,2}^k, v_{\ell,3}^k\}_{\ell \in \{1,2\}, k \in [81]}$, $\bar{\mathcal{V}} = \{\bar{v}_{\ell,1}^k, \bar{v}_{\ell,2}^k, \bar{v}_{\ell,3}^k\}_{\ell \in \{1,2\}, k \in [81]}$, $\mathcal{Z} = \{z_1, \dots, z_{729}\}$. The budget of extra capacities is $B = 486$, and the objective value is $\bar{K} = 2187 + 729 + 1944 + 729 - 162 = 5427$. The preferences of the agents in the new instance are as follows.

$$\begin{array}{ll}
a_1 : z_1, \dots & z_1 : a_1, \dots \\
\vdots & \vdots \\
a_{729} : z_{729}, \dots & z_{729} : a_{729}, \dots \\
& x_1 : \dots \\
& \vdots \\
& x_{2187} : \dots \\
\forall k \in [81] : & \\
i_3^k : r_{i_3^k}^k & j_1^k : w_{1,1}^k, w_{2,2}^k, i_3^k, \dots \\
& j_2^k : w_{2,3}^k, w_{1,2}^k, i_3^k, \dots \\
& j_3^k : w_{1,3}^k, i_3^k, w_{2,1}^k, \dots \\
\text{residents in } \mathcal{B}_1^k & \text{hospitals in } \mathcal{B}_1^k \\
\text{residents in } \mathcal{B}_2^k & \text{hospitals in } \mathcal{B}_2^k
\end{array}$$

where the preference list of i_3^k is $r_{i_3^k}^k = v_1^{k+1}, v_2^{k+1}, v_3^{k+1}, \dots, v_2^{k+4}, j_2^k, v_3^{k+4}, \dots, v_1^{k+7}, j_1^k, v_2^{k+7}, \dots, v_3^{k+9}, j_3^k, \mathcal{Z}, \dots, \mathcal{X}$.

To illustrate, we present the preference lists of the residents in $\mathcal{B}_1^k$ and the preference lists of the hospitals in $\mathcal{B}_1^k$. Given $k \in [81]$, the preference lists of the residents in $\mathcal{B}_1^k$ are as follows.

$$\begin{array}{l}
w_{1,1}^k : v_{1,1}^k, \bar{v}_{1,2}^k, \bar{v}_{1,3}^k, v_{1,1}^{k+1}, \dots, v_{1,3}^{k+3}, j_1^k, \mathcal{Z}, \dots, \mathcal{X} \\
w_{1,2}^k : v_{1,2}^k, \bar{v}_{1,1}^k, \bar{v}_{1,3}^k, v_{1,1}^{k+1}, \dots, v_{1,1}^{k+4}, j_2^k, \mathcal{Z}, \dots, \mathcal{X} \\
w_{1,3}^k : v_{1,3}^k, \bar{v}_{1,1}^k, \bar{v}_{1,2}^k, v_{1,1}^{k+1}, \dots, v_{1,3}^{k+9}, j_3^k, \mathcal{Z}, \dots, \mathcal{X} \\
y_1^k : \bar{v}_{1,2}^k, \bar{v}_{1,1}^k, \bar{v}_{1,3}^k, v_{1,1}^{k+1}, \mathcal{Z}, \dots, \mathcal{X} \\
u_{1,1}^k : v_{1,1}^k, \dots \\
u_{1,2}^k : v_{1,2}^k, \dots \\
u_{1,3}^k : v_{1,3}^k, \dots \\
\bar{u}_{1,1}^k : \bar{v}_{1,1}^k, \dots \\
\bar{u}_{1,2}^k : \bar{v}_{1,2}^k, \dots \\
\bar{u}_{1,3}^k : \bar{v}_{1,3}^k, \dots
\end{array}$$

Given $k \in [81]$, the preference lists of the hospitals in $\mathcal{B}_1^k$ are as follows.

$$\begin{array}{l}
v_{1,1}^k : u_{1,1}^k, \mathcal{A}, w_{1,1}^k, \{w_{1,1}^{k'}\}_{k' \neq k}, \{\mathcal{W}_\ell^k\}_{\ell, k} \setminus \{w_{1,1}^{k'}\}_{k'}, \dots, \mathcal{Y} \\
v_{1,2}^k : u_{1,2}^k, \mathcal{A}, w_{1,2}^k, \{w_{1,2}^{k'}\}_{k' \neq k}, \{\mathcal{W}_\ell^k\}_{\ell, k} \setminus \{w_{1,2}^{k'}\}_{k'}, \dots, \mathcal{Y} \\
v_{1,3}^k : u_{1,3}^k, \mathcal{A}, w_{1,3}^k, \{w_{1,3}^{k'}\}_{k' \neq k}, \{\mathcal{W}_\ell^k\}_{\ell, k} \setminus \{w_{1,3}^{k'}\}_{k'}, \dots, \mathcal{Y} \\
\bar{v}_{1,1}^k : \bar{u}_{1,1}^k, \mathcal{A}, \mathcal{W}_1^k \setminus \{w_{1,1}^k\}, y_1^k, \{\mathcal{W}_1^{k'}\}_{1, k' \neq k} \setminus \{w_{1,1}^{k'}\}_{k' \neq k}, \dots, \mathcal{Y} \setminus \{y_1^k\} \\
\bar{v}_{1,2}^k : \bar{u}_{1,2}^k, \mathcal{A}, \mathcal{W}_1^k \setminus \{w_{1,2}^k\}, y_1^k, \{\mathcal{W}_1^{k'}\}_{1, k' \neq k} \setminus \{w_{1,2}^{k'}\}_{k' \neq k}, \dots, \mathcal{Y} \setminus \{y_1^k\} \\
\bar{v}_{1,3}^k : \bar{u}_{1,3}^k, \mathcal{A}, \mathcal{W}_1^k \setminus \{w_{1,3}^k\}, y_1^k, \{\mathcal{W}_1^{k'}\}_{1, k' \neq k} \setminus \{w_{1,3}^{k'}\}_{k' \neq k}, \dots, \mathcal{Y} \setminus \{y_1^k\}
\end{array}$$

For $k \in [81]$, the preference lists of the residents in $\mathcal{B}_2^k$ and the preference lists of the hospitals in $\mathcal{B}_2^k$ are similar.

Given a weakly stable matching in the instance of MIN-w SMT, we find a stable matching in the reduced instance of MIN-AVG_{EXP} HR.

We consider the weakly stable matching $\mu = \{(i_1, j_1), (i_2, j_3), (i_3, j_2)\}$, which has cost 3. The reduction allocates one extra capacity to each school $v_{1,2}^k, v_{1,3}^k, \bar{v}_{1,1}^k$ of village $\mathcal{B}_1^k$, and one extra capacity to each school $v_{2,2}^k, v_{2,3}^k, \bar{v}_{2,1}^k$ of village $\mathcal{B}_2^k$, for every $k \in [81]$. Note that a total of 486 extra capacities are allocated, thus exhausting the overall budget B . Therefore, the matching induced by the reduction is made of the following pairs: $\{(a_i, z_i)\}_{i \in [729]}$, and, for every $k \in [81]$, $\{(w_{\ell,2}^k, v_{\ell,2}^k), (w_{\ell,3}^k, v_{\ell,3}^k), (y_\ell^k, \bar{v}_{\ell,1}^k)\}_{\ell \in \{1,2\}}$, $(w_{1,1}^k, j_1^k)$, $(w_{2,1}^k, j_3^k)$, (i_3^k, j_2^k) , $\{(u_{\ell,h}^k, v_{\ell,h}^k), (\bar{u}_{\ell,h}^k, \bar{v}_{\ell,h}^k)\}_{\ell \in [2], h \in [3]}$. The overall cost of the matching just described is $729 + 81 \cdot (2 \cdot (1 + 1 + 2) + 13 + 13 + 12 + 2 \cdot 3 \cdot (1 + 1)) = 5427$, which is exactly equal to the objective value $\bar{K}$ of the reduced instance. In Figure A.2 we provide a graphic illustration of the matching in the reduced instance. Note that $w_{1,1}^k$ ($w_{2,1}^k$) is the only resident of village $\mathcal{B}_1^k$ ($\mathcal{B}_2^k$) that is not matched with a school in their own village.

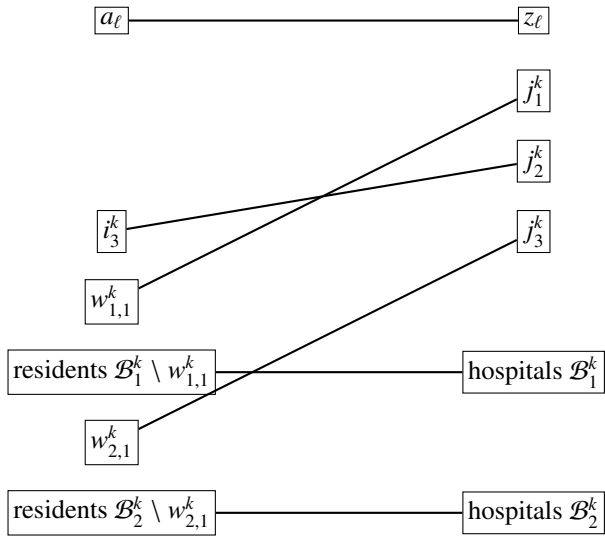


Figure A.2: Example of stable matching in the reduced instance. Note: $\ell \in [729]$ and $k \in [81]$.