Matchings under Stability, Minimum Regret, and Forced and Forbidden Pairs in Marriage Problem

Mandal, Pinaki and Roy, Souvik

18 April 2021

Matchings under Stability, Minimum Regret, and Forced and Forbidden Pairs in Marriage Problem

Pinaki Mandal* and Souvik Roy†

Abstract

We provide a class of algorithms, called men-women proposing deferred acceptance (MWPDA) algorithms, that can produce all stable matchings at every preference profile for the marriage problem. Next, we provide an algorithm that produces a minimum regret stable matching at every preference profile. We also show that its outcome is always women-optimal in the set of all minimum regret stable matchings. Finally, we provide an algorithm that produces a stable matching with given sets of forced and forbidden pairs at every preference profile, whenever such a matching exists. As before, here too we show that the outcome of the said algorithm is women-optimal in the set of all stable matchings with given sets of forced and forbidden pairs.

Keywords: Two-sided matching; Marriage problem; Pairwise stability; Stability; Minimum regret; Forced and forbidden pairs

JEL Classification: C78

*Economic Research Unit, Indian Statistical Institute, Kolkata 700108, India. Email: pnk.rana@gmail.com

†Economic Research Unit, Indian Statistical Institute, Kolkata 700108, India. Email: souvik.2004@gmail.com

1 Introduction

This paper explores the possibilities of designing mechanisms satisfying properties such as (pairwise) stability, minimum regret, and forced and forbidden pairs in case of two-sided one-to-one matching problem (marriage problem).

(Pairwise) stability is a well-known property of a matching. Gale and Shapley (1962) provide an algorithm called *men-proposing/women-proposing deferred acceptance (MPDA/WPDA)* algorithm that produces a stable matching at every preference profile. It is well-known that the outcome of the MPDA (WPDA) algorithm is (i) men-maximal (women-maximal), that is, such an outcome maximizes the match of each man (woman) over all stable matchings, and (ii) women-pessimal (men-pessimal), that is, such an outcome minimizes the match of each woman (man) over all stable matchings.[1]

The main motivation of this paper is to provide an algorithmic characterization of all stable matchings at every preference profile. The other motivation is to provide algorithms to construct stable matchings with additional desirable properties such as minimum regret and forced/forbidden pairs. The importance of a characterization of all stable matchings is well-established in the literature. McVitie and Wilson (1971) provide an iterative procedure to compute all stable matchings for the marriage problem and Martınez et al. (2004) extend that algorithm to two-sided many-to-many matching problem with *substitutable* preferences.[2] Irving and Leather (1986) provide an alternative method of computing all stable matchings for the marriage problem by using the lattice structure of the set of stable matchings. To the best of our knowledge, apart from *Gale-Shapley algorithm*, no direct algorithm that produces stable matching is introduced to the literature.[3] However, as discussed earlier, stable matchings produced by Gale-Shapley algorithm (Gale and Shapley, 1962) suffer from the problem that they are either extremely biased against men (in case of WPDA algorithm) or that towards women (in case of MPDA algorithm).

We present a class of algorithms that we call *men-women proposing deferred acceptance (MWPDA)* algorithms which can produce all stable matchings at every preference profile. Such an algorithm is based on a given collection of cut-off parameters one for each man. A cut-off parameter κ_m for a man m is an arbitrary integer between 1 and the number of women plus one. For a given collection of cut-off parameters the algorithm works in a sequence of stages as follows. At the beginning of Stage 1, each man m proposes each acceptable woman who appears in top κ_m positions according to his preference, and then WPDA algorithm is performed with respect to the proposals that the women receive. From a given stage we go to the subsequent stage if there is a man who (i) has not yet proposed all acceptable women according to

[1] See Gale and Shapley (1962), McVitie and Wilson (1971), Knuth (1976), and Abdulkadiroglu and Sönmez (2013) for details.

[2] Kelso Jr and Crawford (1982) are the first to use the substitutability property to show the existence of stable matchings in a many-to-one model with money.

[3] McVitie and Wilson (1971) provide a method to compute all stable matchings at a preference profile. However, their method is lengthy in the sense that every time one needs to produce some particular stable matching, he/she has to start from the men-maximal (or women-maximal) stable matching and keep on producing all stable matchings that come in the process before he/she arrives at the intended stable matching. Another problem with this method is that it is not structured enough to produce stable matching with additional desirable properties.

his preference, and (ii) is unmatched at that given stage. Moreover, in any stage, if a man m was matched in the previous stage, then he proposes the same set of women as he did in the previous stage, otherwise he proposes the remaining set of acceptable women (that is, the acceptable women who do not appear in top κ_m positions according to his preference).

Theorem 3.1 of our paper shows that the outcome of an MWPDA algorithm is stable at every preference profile for any cut-off vector. Theorem 3.2 shows that for any stable matching at a preference profile, there is a cut-off vector such that the MWPDA algorithm with respect to it will produce that stable matching. Theorem 3.3 provides a necessary and sufficient condition on the cut-off vectors so that the MWPDA algorithms with those cut-off vectors will converge at the first stage. We also discuss that these algorithms can be extended to produce all stable matchings in a two-sided many-to-one matching problem (college admissions problem) in a way mentioned in Roth and Sotomayor (1989).

The notion of *minimum regret under stability* is introduced in Knuth (1976). It captures the idea of a Rawlsian welfare function. The regret of an agent in a matching is defined as the rank of his/her match according to his/her preference, and the regret of a matching is defined as the highest regret (over all agents) at that matching. A stable matching satisfies *minimum regret stable* property at a preference profile if it has the minimum regret among all the stable matchings at that preference profile.[4] Both MPDA and WPDA algorithms are far from satisfying the minimum regret under stability as their outcomes are either women-pessimal or men-pessimal. We provide a direct algorithm called the *sequential MWPDA* algorithm that produces a minimum regret stable matching at every preference profile.[5] We further show that the outcome of the sequential MWPDA algorithm is women-optimal in the set of all minimum regret stable matchings.

For practical reasons, sometimes one needs to construct stable matching with additional constraints. The notion of stable matching with *forced pairs* is introduced in Knuth (1976), and that with *forbidden pairs* is introduced in Dias et al. (2003). To the best of our knowledge, there is no direct algorithm that produces stable matching with these properties.[6] We provide an algorithm called the *conditional MWPDA algorithm* that produces stable matching with given sets of forced and forbidden pairs, whenever such a matching exists. We further show that whenever the conditional MWPDA algorithm produces such a matching, the outcome is women-optimal in the set of all stable matchings with given sets of forced and forbidden pairs.

[4]Note that the regret of an unstable matching can be strictly less than the minimum regret under stability.

[5]Knuth (1976) provides an algorithm with runtime of the order $O(n^4)$ to find a minimum regret stable matching where n is the number of men, as well as the number of women. The algorithm given in Knuth (1976) is attributed to Alan Selkow. Later, Gusfield (1987) provide an algorithm that terminates in $O(n^2)$ time.

[6]Knuth (1976) provides an algorithm that produces a stable matching with a given set of forced pairs or reports that none exists, in $O(n^2)$ time, where n is the number of men, as well as the number of women. Later, Gusfield and Irving (1989) provide an algorithm that terminates in $O(|Q_1|^2)$ time, after pre-processing the preference lists in $O(n^4)$ time, where Q_1 is the set of given forced pairs. Dias et al. (2003) provide a computer algorithm that produces a stable matching with a given set of forced pairs Q_1 and a given set of forbidden pairs Q_2 in $O((|Q_1| + |Q_2|)^2)$ time, after pre-processing the preference lists in $O(n^4)$ time.

1.1 Organization of the paper

The paper is organized as follows. The marriage problem framework is presented in Section 2. In Section 3, we present MWPDA algorithms and show that they produce all stable matchings at every preference profile for the marriage problem. We also provide a necessary and sufficient condition for the convergence of these algorithms at the first stage, and discuss how these algorithms can be used to construct all stable matchings for the college admissions problem. In Section 4, we present an algorithm that produces a minimum regret stable matching at every preference profile, and in Section 5, we present an algorithm that produces a stable matching with given sets of forced and forbidden pairs. All the proofs are collected in the Appendix.

2 Model

For a finite set A, let $\mathbb{L}(A)$ denote the set of all strict linear orders over A.[7] An element P of $\mathbb{L}(A)$ is called a ***preference*** over A. For a preference $P \in \mathbb{L}(A)$, let R denote the weak part of P, that is, for all $a, b \in A$, aRb if and only if $[aPb \text{ or } a = b]$.

For $P \in \mathbb{L}(A)$ and $1 \leq k \leq |A|$, we define $T_k(P) := \{b \in A : |\{a : aRb\}| \leq k\}$. So, $T_k(P)$ is the set of top k elements of A according to P. Moreover, for $P \in \mathbb{L}(A)$ and $a \in A$, we define $rank(P, a) = k$ if $\left|\{b \in A : bPa\}\right| = k - 1$.

We introduce a specialized model of the two-sided matching problem, which will turn out to be sufficiently general to explore the general problem. The simplest two-sided matching problem to model is the "marriage problem", which consists of two (finite) sets of agents $M = \{m_1, \ldots, m_p\}$ and $W = \{w_1, \ldots, w_q\}$ ("men" and "women"). Throughout this paper, we assume $p, q \geq 2$. We denote by $N = M \cup W$. Each $m \in M$ has a preference $P_m \in \mathbb{L}(W \cup \{\emptyset\})$ and each $w \in W$ has a preference $P_w \in \mathbb{L}(M \cup \{\emptyset\})$. A man m (woman w) is called ***acceptable*** for a woman w (man m) at a preference P_w (P_m) if $mP_w\emptyset$ ($wP_m\emptyset$). For $m \in M$ ($w \in W$), we denote by $\mathcal{A}(P_m)$ ($\mathcal{A}(P_w)$) the set of acceptable women (men) for m (w) at a preference P_m (P_w). By $P_N = (P_{m_1}, \ldots, P_{m_p}, P_{w_1}, \ldots P_{w_q})$, we denote a vector of all the agents' preferences, which will be referred to as a ***preference profile***.

Definition 2.1. A ***matching*** between M and W is a function $\mu : N \to N \cup \{\emptyset\}$ such that

(i) $\mu(m) \in W \cup \{\emptyset\}$ for all $m \in M$,

(ii) $\mu(w) \in M \cup \{\emptyset\}$ for all $w \in W$, and

(iii) $\mu(m) = w$ if and only if $\mu(w) = m$.

Definition 2.2. A matching $\mu : N \to N \cup \{\emptyset\}$ is ***individually rational*** at a preference profile P_N if $\mu(a)R_a\emptyset$ for all $a \in N$.

[7] A ***strict linear order*** is a semiconnex, asymmetric, and transitive binary relation.

Definition 2.3. A pair $(m, w) \in M \times W$ is called a *blocking pair* of a matching $\mu : N \to N \cup \{\emptyset\}$ at a preference profile P_N if $wP_m\mu(m)$ and $mP_w\mu(w)$.

A matching $\mu : N \to N \cup \{\emptyset\}$ is called ***pairwise stable*** at a preference profile P_N if it is individually rational and has no blocking pairs at P_N.

Definition 2.4. A coalition $N' \subseteq N$ is called a *blocking coalition* of a matching $\mu : N \to N \cup \{\emptyset\}$ at a preference profile P_N if there exists another matching $\mu' : N \to N \cup \{\emptyset\}$ such that

(i) $\mu'(a) \in N' \cup \{\emptyset\}$ for all $a \in N'$, and

(ii) $\mu'(a)P_a\mu(a)$ for all $a \in N'$.

If a matching $\mu : N \to N \cup \{\emptyset\}$ has no blocking coalition at a preference profile P_N, then it is called ***stable*** at P_N.

Remark 2.1. It is well-known that pairwise stability and stability are equivalent.[8] For this reason, we will say a matching is stable at a preference profile if and only if it is pairwise stable at that preference profile.

We denote by $\mathcal{C}(P_N)$ the set of all stable matchings at a preference profile P_N. It is well-known that $\mathcal{C}(P_N) \neq \emptyset$ for every preference profile P_N (see Gale and Shapley (1962) for details).

Definition 2.5. For a preference profile P_N and a set of matchings $\mathcal{M}$, a matching $\mu \in \mathcal{M}$ is ***women-optimal in*** $\mathcal{M}$ at P_N if $\mu(w)R_w\mu'(w)$ for all $w \in W$ and all $\mu' \in \mathcal{M}$. Similarly, one can define the notion a men-optimal matching in a set of matchings.[9]

A matching $\mu \in \mathcal{C}(P_N)$ is ***men-optimal (women-optimal) stable matching*** at P_N if μ is men-optimal (women-optimal) in $\mathcal{C}(P_N)$ at P_N.

It is well-known that a men-optimal (women-optimal) stable matching exists at every preference profile (see Gale and Shapley (1962) for details).

3 Algorithms for producing all stable matchings at a preference profile

An ***algorithm*** is a procedure that produces a matching at any preference profile. In this section, we provide a class of algorithms, called ***men-women proposing deferred acceptance (MWPDA)*** algorithms, which can produce every stable matching at a preference profile. These algorithms are built on well-known *deferred acceptance (DA)* algorithms. For the sake of completeness, we begin with a description (that is suitable for our purpose) of DA algorithms.

[8] See Roth and Sotomayor (1992) for details.

[9] Women-optimal (men-optimal) matching in an arbitrary set of matchings may not exist.

3.1 Deferred Acceptance algorithm

There are two types of deferred acceptance algorithms: women-proposing deferred acceptance (WPDA) and men-proposing deferred acceptance (MPDA). In the following, we provide a description of the WPDA algorithm at a preference profile P_N. The same of the MPDA algorithm can be obtained by interchanging the roles of women and men in the WPDA algorithm.

Step 1. Every woman w proposes her top-ranked acceptable man according to P_w[10]. Then, every man m who has at least one proposal keeps (tentatively) the top acceptable woman according to P_m among these proposals and rejects the rest. Denote the tentative matching thus obtained by μ_1.

Step 2. Every woman w who was rejected in the previous step, proposes the top acceptable man among those men who have not rejected her in earlier steps. Then, every man m who has at least one proposal, including any proposal tentatively kept from earlier steps, keeps (tentatively) the top acceptable woman among these proposals and rejects the rest. Denote the tentative matching thus obtained by μ_2.

$$\vdots$$

The process is then repeated from Step 2 till a step such that for each woman one of the following two happens: (i) she has proposed all acceptable men, (ii) she is accepted by some man who is acceptable to her. At this point, the tentative proposal accepted by a man becomes permanent. Call this the outcome of the WPDA algorithm at P_N.

Remark 3.1. Gale and Shapley (1962) show that at every preference profile P_N, there exists a unique men-optimal stable matching that is produced by the MPDA algorithm and a unique women-optimal stable matching that is produced by the WPDA algorithm.

Throughout this paper, we denote the men-optimal and the women-optimal stable matching at a preference profile P_N by $\mu_M(P_N)$ and $\mu_W(P_N)$, respectively. Moreover, whenever the preference profile P_N is clear from the context, we drop it from these notations, that is, we write μ_M for $\mu_M(P_N)$, etc.

Remark 3.2. For all $\mu \in \mathcal{C}(P_N)$, $\mu_M(m) R_m \mu(m) R_m \mu_W(m)$ for all $m \in M$, and $\mu_W(w) R_w \mu(w) R_w \mu_M(w)$ for all $w \in W$.[11]

3.2 MWPDA algorithms

We begin with introducing a piece of notation that will simplify the presentation of our algorithm. For a preference $P_w \in \mathbb{L}(M \cup \{\emptyset\})$ and $M' \subseteq M$, define $P_w^{M'}$ as the preference that is obtained by moving the elements of $M' \cup \{\emptyset\}$ to the top of P_w maintaining their relative ordering. More formally, $P_w^{M'}$ is such that

[10] That is, if the top-ranked man of a woman is acceptable, then she proposes him, otherwise she does not propose anybody.

[11] See Gale and Shapley (1962), McVitie and Wilson (1971), Knuth (1976), and Abdulkadiroglu and Sönmez (2013) for details.

(i) for all $x, y \in M' \cup \{\emptyset\}$, $x P_w^{M'} y$ if and only if $x P_w y$, and (ii) for all $x \in M' \cup \{\emptyset\}$ and $y \in M \setminus M'$, we have $x P_w^{M'} y$.[12]

An MWPDA algorithm is parameterized by a *cut-off vector*. A ***cut-off vector*** is defined as $\kappa = (\kappa_{m_1}, \ldots, \kappa_{m_p})$, where for all $m \in M$, $\kappa_m \in \{1, \ldots, q+1\}$ is the cut-off parameter of man m. An MWPDA algorithm involves a sequence of stages. At the beginning of a stage, say Stage s, each man m proposes a set of women (which is determined by the parameters). We denote this set by $W^s(m)$. The set of proposals that each $w \in W$ receives in that stage is denoted by $M^s(w)$, that is, $M^s(w) = \{m : w \in W^s(m)\}$.

Below, we present a detailed description (using the notations introduced above) of the MWPDA algorithm with cut-off vector κ at a preference profile P_N.

Stage 1. Take $W^1(m) = T_{\kappa_m}(P_m) \cap \mathcal{A}(P_m)$ for all $m \in M$. Perform the WPDA algorithm at the preference profile $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Let μ^1 be the outcome. If $W^1(m) = \mathcal{A}(P_m)$ for all $m \in M$ with $\mu^1(m) = \emptyset$, then conclude that the algorithm converges and define μ^1 as the outcome of the algorithm. Otherwise, go to Stage 2.

Stage 2. For all $m \in M$, take $W^2(m)$ such that

$$W^2(m) = \begin{cases} W^1(m) & \text{if } \mu^1(m) \neq \emptyset; \\ \mathcal{A}(P_m) \setminus W^1(m) & \text{if } \mu^1(m) = \emptyset \text{ and } W^1(m) \subsetneq \mathcal{A}(P_m); \\ \emptyset & \text{if } \mu^1(m) = \emptyset \text{ and } W^1(m) = \mathcal{A}(P_m).\text{[13]} \end{cases}$$

Perform the WPDA algorithm at the preference profile $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^2(w_1)}, \ldots, P_{w_q}^{M^2(w_q)})$. Let μ^2 be the outcome. If $W^1(m) \cup W^2(m) = \mathcal{A}(P_m)$ for all $m \in M$ with $\mu^2(m) = \emptyset$, then conclude that the algorithm converges and define μ^2 as the outcome of the algorithm. Otherwise, go to Stage 3.

Stage 3. For all $m \in M$, take $W^3(m)$ such that

$$W^3(m) = \begin{cases} W^2(m) & \text{if } \mu^2(m) \neq \emptyset; \\ \mathcal{A}(P_m) \setminus \big(\bigcup\limits_{s \leq 2} W^s(m) \big) & \text{if } \mu^2(m) = \emptyset \text{ and } \bigcup\limits_{s \leq 2} W^s(m) \subsetneq \mathcal{A}(P_m); \\ \emptyset & \text{if } \mu^2(m) = \emptyset \text{ and } \bigcup\limits_{s \leq 2} W^s(m) = \mathcal{A}(P_m). \end{cases}$$

Perform the WPDA algorithm at the preference profile $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^3(w_1)}, \ldots, P_{w_q}^{M^3(w_q)})$. Let μ^3 be the outcome. If $\bigcup\limits_{s \leq 3} W^s(m) = \mathcal{A}(P_m)$ for all $m \in M$ with $\mu^3(m) = \emptyset$, then conclude that the algorithm converges and define μ^3 as the outcome of the algorithm. Otherwise, go to Stage 4.

[12] Note that such a preference $P_w^{M'}$ may not be unique since it does not specify the relative ranking of the elements of $M \setminus M'$.

[13] It follows from the definition of $W^1(m)$ that $W^1(m) \subseteq \mathcal{A}(P_m)$ for all $m \in M$. Therefore, the cases considered in this definition are exhaustive.

$$\vdots$$

We continue this till a stage t^* such that $\bigcup\limits_{s \leq t^*} W^s(m) = \mathcal{A}(P_m)$ for all $m \in M$ with $\mu^{t^*}(m) = \emptyset$. Since both the number of men and the number of women are finite, such a stage t^* must exist. At this stage, define the matching μ^{t^*} as the outcome the algorithm.

Remark 3.3. If $\kappa_m = q + 1$ for all $m \in M$, then the MWPDA algorithm with κ boils down to the WPDA algorithm.

We illustrate MWPDA algorithm by means of the following example.

Example 3.1. Let $M = \{m_1, m_2, m_3, m_4, m_5\}$ and $W = \{w_1, w_2, w_3, w_4, w_5\}$. Consider the preference profile P_N as given below:

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{m_5}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	P_{w_5}
w_1	w_1	w_2	w_1	w_1	m_2	m_4	m_5	m_2	m_3
w_2	w_3	w_1	w_2	w_2	m_5	m_5	m_2	m_3	m_1
w_3	w_2	w_3	w_5	w_3	m_1	m_2	m_4	m_1	m_5
w_4	w_4	w_4	w_4	w_4	$\emptyset$	m_1	m_3	m_5	$\emptyset$
w_5	w_5	w_5	w_3	w_5	m_3	m_3	$\emptyset$	m_4	m_2
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	m_4	$\emptyset$	m_1	$\emptyset$	m_4

Table 3.1: Preference profile for Example 3.1

Let the cut-off vector κ be such that $\kappa_{m_1} = 2, \kappa_{m_2} = 4, \kappa_{m_3} = 3, \kappa_{m_4} = 1$ and $\kappa_{m_5} = 2$. The MWPDA algorithm with κ at the preference profile P_N given in Table 3.1 works as follows.

Stage 1. Perform the WPDA algorithm at the preference profile $(P_{m_1}, \ldots, P_{m_5}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_5}^{M^1(w_5)})$ given in Table 3.2. The dots in Table 3.2 indicate that all preferences for the corresponding parts are irrelevant and can be chosen arbitrarily. To emphasize the process at Stage 1, for each man m we have highlighted the women in P_m in blue that m proposes, and for each woman w we have highlighted the men in P_w in blue who propose her.

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{m_5}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	P_{w_5}	$P_{w_1}^{M^1(w_1)}$	$P_{w_2}^{M^1(w_2)}$	$P_{w_3}^{M^1(w_3)}$	$P_{w_4}^{M^1(w_4)}$	$P_{w_5}^{M^1(w_5)}$
w_1	w_1	w_2	w_1	w_1	m_2	m_4	m_5	m_2	m_3	m_2	m_5	m_2	m_2	$\emptyset$
w_2	w_3	w_1	w_2	w_2	m_5	m_5	m_2	m_3	m_1	m_5	m_2	m_3	$\emptyset$	$\vdots$
w_3	w_2	w_3	w_5	w_3	m_1	m_2	m_4	m_1	m_5	m_1	m_1	$\emptyset$	$\vdots$	
w_4	w_4	w_4	w_4	w_4	$\emptyset$	m_1	m_3	m_5	$\emptyset$	$\emptyset$	m_3	$\vdots$		
w_5	w_5	w_5	w_3	w_5	m_3	m_3	$\emptyset$	m_4	m_2	m_3	$\emptyset$			
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	m_4	$\emptyset$	m_1	$\emptyset$	m_4	m_4	$\vdots$			

Table 3.2: Updated preference profile at Stage 1

The outcome of the WPDA algorithm at Stage 1 is $[(m_1,\emptyset),(m_2,w_1),(m_3,w_3),(m_4,\emptyset),(m_5,w_2)]$. Since $\mu^1(m_1)=\emptyset$ with $W^1(m_1)\subsetneq \mathcal{A}(P_{m_1})$, we go to Stage 2.

Stage 2. Perform the WPDA algorithm at the preference profile $(P_{m_1},\dots,P_{m_5},P_{w_1}^{M^2(w_1)},\dots,P_{w_5}^{M^2(w_5)})$ given in Table 3.3.

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{m_5}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	P_{w_5}	$P_{w_1}^{M^2(w_1)}$	$P_{w_2}^{M^2(w_2)}$	$P_{w_3}^{M^2(w_3)}$	$P_{w_4}^{M^2(w_4)}$	$P_{w_5}^{M^2(w_5)}$
w_1	w_1	w_2	w_1	w_1	m_2	m_4	m_5	m_2	m_3	m_2	m_4	m_2	m_2	m_1
w_2	w_3	w_1	w_2	w_2	m_5	m_5	m_2	m_3	m_1	m_5	m_5	m_4	m_1	$\emptyset$
w_3	w_2	w_3	w_5	w_3	m_1	m_2	m_4	m_1	m_5	$\emptyset$	m_2	m_3	m_4	m_4
w_4	w_4	w_4	w_4	w_4	$\emptyset$	m_1	m_3	m_5	$\emptyset$	m_3	m_3	$\emptyset$	$\emptyset$	$\vdots$
w_5	w_5	w_5	w_3	w_5	m_3	m_3	$\emptyset$	m_4	m_2	$\vdots$	$\emptyset$	m_1	$\vdots$	
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	m_4	$\emptyset$	m_1	$\emptyset$	m_4		$\vdots$	$\vdots$		

Table 3.3: Updated preference profile at Stage 2

The outcome of the WPDA algorithm at Stage 2 is $[(m_1,w_4),(m_2,w_1),(m_3,w_3),(m_4,w_2),(m_5,\emptyset)]$. Since $\mu^2(m_5)=\emptyset$ with $W^1(m_5)\cup W^2(m_5)\subsetneq \mathcal{A}(P_{m_5})$, we go to Stage 3.

Stage 3. Perform the WPDA algorithm at the preference profile $(P_{m_1},\dots,P_{m_5},P_{w_1}^{M^3(w_1)},\dots,P_{w_5}^{M^3(w_5)})$ given in Table 3.4.

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{m_5}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	P_{w_5}	$P_{w_1}^{M^3(w_1)}$	$P_{w_2}^{M^3(w_2)}$	$P_{w_3}^{M^3(w_3)}$	$P_{w_4}^{M^3(w_4)}$	$P_{w_5}^{M^3(w_5)}$
w_1	w_1	w_2	w_1	w_1	m_2	m_4	m_5	m_2	m_3	m_2	m_4	m_5	m_2	m_1
w_2	w_3	w_1	w_2	w_2	m_5	m_5	m_2	m_3	m_1	$\emptyset$	m_2	m_2	m_1	m_5
w_3	w_2	w_3	w_5	w_3	m_1	m_2	m_4	m_1	m_5	m_3	m_3	m_4	m_5	$\emptyset$
w_4	w_4	w_4	w_4	w_4	$\emptyset$	m_1	m_3	m_5	$\emptyset$	$\vdots$	$\emptyset$	m_3	m_4	m_4
w_5	w_5	w_5	w_3	w_5	m_3	m_3	$\emptyset$	m_4	m_2		$\vdots$	$\emptyset$	$\emptyset$	$\vdots$
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	m_4	$\emptyset$	m_1	$\emptyset$	m_4			m_1	$\vdots$	

Table 3.4: Updated preference profile at Stage 3

The outcome of the WPDA algorithm at Stage 3 is $[(m_1, w_4), (m_2, w_1), (m_3, \emptyset), (m_4, w_2), (m_5, w_3)]$. Since $\mu^3(m_3) = \emptyset$ with $W^1(m_3) \cup W^2(m_3) \cup W^3(m_3) \subsetneq \mathcal{A}(P_{m_3})$, we go to Stage 4.

Stage 4. Perform the WPDA algorithm at the preference profile $(P_{m_1}, \ldots, P_{m_5}, P_{w_1}^{M^4(w_1)}, \ldots, P_{w_5}^{M^4(w_5)})$ given in Table 3.5.

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{m_5}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	P_{w_5}	$P_{w_1}^{M^3(w_1)}$	$P_{w_2}^{M^3(w_2)}$	$P_{w_3}^{M^3(w_3)}$	$P_{w_4}^{M^3(w_4)}$	$P_{w_5}^{M^3(w_5)}$
w_1	w_1	w_2	w_1	w_1	m_2	m_4	m_5	m_2	m_3	m_2	m_4	m_5	m_2	m_3
w_2	w_3	w_1	w_2	w_2	m_5	m_5	m_2	m_3	m_1	$\emptyset$	m_2	m_2	m_3	m_1
w_3	w_2	w_3	w_5	w_3	m_1	m_2	m_4	m_1	m_5	$\vdots$	$\emptyset$	m_4	m_1	m_5
w_4	w_4	w_4	w_4	w_4	$\emptyset$	m_1	m_3	m_5	$\emptyset$		$\vdots$	$\emptyset$	m_5	$\emptyset$
w_5	w_5	w_5	w_3	w_5	m_3	m_3	$\emptyset$	m_4	m_2			m_1	m_4	m_4
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	m_4	$\emptyset$	m_1	$\emptyset$	m_4			$\vdots$	$\emptyset$	$\vdots$

Table 3.5: Updated preference profile at Stage 4

The outcome of the WPDA algorithm at Stage 4 is $[(m_1, w_5), (m_2, w_1), (m_3, w_4), (m_4, w_2), (m_5, w_3)]$. Since $\mu^4(m) \neq \emptyset$ for all $m \in M$, the outcome of MWPDA algorithm with the cut-off vector κ is $[(m_1, w_5), (m_2, w_1), (m_3, w_4), (m_4, w_2), (m_5, w_3)]$.

3.3 MWPDA algorithms produce all stable matchings

In this subsection, we explore the stability of the outcome of an MWPDA algorithm. We also provide a sufficient condition on an MWPDA algorithm to produce a specific stable matching at the first step of the WPDA algorithm at Stage 1 of the mentioned MWPDA algorithm. Our next theorem shows that the outcome of an MWPDA algorithm at any preference profile and with any cut-off vector is stable.

Theorem 3.1. *For every preference profile P_N and every cut-off vector κ, the MWPDA algorithm with κ produces a stable matching at P_N.*

The proof of this theorem is relegated to Appendix A; here we provide the idea of it. By Observation A.1, the match of each man (weakly) improves (according to his preference) over the steps of the WPDA algorithm at any given stage. Next, we show the match of each woman (weakly) improves over the stages (Lemma A.1). Finally, we combine these two facts to prove Theorem 3.1.

Now, we present the main result of this section. It says that every stable matching at any preference profile can be produced by an MWPDA algorithm with some cut-off vector. However, we prove a stronger version of this, which says that every stable matching at a preference profile can be produced at the first step of the WPDA algorithm at Stage 1 of an MWPDA algorithm by using a *suitable* cut-off vector.

Theorem 3.2. *Let P_N be a preference profile and let $\mu \in \mathcal{C}(P_N)$. Suppose the cut-off vector κ is such that $\kappa_m = rank(P_m, \mu(m))$ for all $m \in M$. Then, the MWPDA algorithm with cut-off vector κ produces μ at P_N. Furthermore, μ is produced at the first step of the WPDA algorithm at Stage 1 (of the mentioned MWPDA algorithm).*

The proof of this theorem is relegated to Appendix B.2. It is worth mentioning that the cut-off vector κ defined in Theorem 3.2 is *not* the unique cut-off vector that produces μ at the first step of the WPDA algorithm at Stage 1.

In view of Theorem 3.2, one may think that if every stable matching can be produced at the first step of the WPDA algorithm at Stage 1 of an MWPDA algorithm, then why do we need a sequence of stages and a sequence of steps of the WPDA algorithm at each stage? The answer to this question is as follows. As it is evident from Theorem 3.2, the 'suitable' cut-off vector for a given stable matching that produces it at the first step of the WPDA algorithm at the first stage *cannot* be identified without using complete knowledge of that stable matching. Thus, in order to find *all* stable matchings at a preference profile, one needs to use MWPDA algorithm with arbitrary cut-off vectors (and consequently needs to go through several stages).

3.4 Convergence of MWPDA algorithms at the first stage

In this subsection, we discuss the convergence of an MWPDA algorithm. As we have mentioned in Subsection 3.3, for every stable matching there exists a cut-off vector so that the MWPDA algorithm with that converges at the first step of the WPDA algorithm at Stage 1 producing the stable matching. However, identifying such a cut-off vector requires complete knowledge of the stable matching. In view of this, we provide a necessary and sufficient condition on the cut-off vectors so that the MWPDA algorithms with those cut-off vectors converge at the first stage.

Recall that, we denote the men-optimal stable matching at a preference profile P_N by $\mu_M(P_N)$. Moreover, whenever the preference profile P_N is clear from the context, we drop it from this notation, that is, we write μ_M for $\mu_M(P_N)$.

Theorem 3.3. *Let P_N be a preference profile. The MWPDA algorithm with a cut-off vector κ at P_N converges at Stage 1 if and only if $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\{|\mathcal{A}(P_m)|, 1\}\right\}$ for all $m \in M$.*

The proof of this theorem is relegated to Appendix B.1.

Remark 3.4. A cut-off vector κ with $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$ does *not* guarantee the convergence of the MWPDA algorithm *at the first step* of the WPDA algorithm at the first stage, it might take several steps to converge.

3.5 Application to the college admissions problem

The "college admissions problem" is a many-to-one generalization of the marriage problem.[14] Every (many-to-one) stable matching in the college admissions problem where colleges' preferences satisfy *responsiveness* can be obtained from Theorem 3.2 in the following way.[15]

(i) Construct a marriage problem for the given college admissions problem (see Roth (1985) and Roth and Sotomayor (1989) for details on how to construct a related marriage problem).

(ii) Apply MWPDA algorithms to obtain all (one-to-one) stable matchings of the marriage problem.

(iii) Transform all (one-to-one) stable matchings of the marriage problem to their many-to-one versions by using a transformation as defined in Roth and Sotomayor (1989).

It follows from Lemma 1 in Roth and Sotomayor (1989) that the many-to-one matchings of the college admissions problem constructed in this manner will be the only pairwise stable matchings, and from Proposition 1 in Roth and Sotomayor (1989), that they will also be the only stable matchings.

4 A minimum regret stable algorithm

In this section, we present an algorithm which produces a stable matching at every preference profile with an additional desirable property, namely minimum regret. As we have mentioned in Remark 3.1, the outcome of the WPDA algorithm is women-optimal stable matching and that of the MPDA algorithm is men-optimal stable matching. In other words, both these algorithms are extremely biased.[16] However, as the following example demonstrates, MWPDA algorithms with suitable cut-off vectors can produce stable matchings that are not so biased.

Example 4.1. Let $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$. Consider the preference profile P_N given in Table 4.1.

[14] See Abdulkadiroglu and Sönmez (2013) for a formal description of the college admissions problem.

[15] The notion of responsiveness is due to Roth (1985), see Abdulkadiroglu and Sönmez (2013) for a formal definition of the same.

[16] See Remark 3.2 for details.

P_{m_1}	P_{m_2}	P_{m_3}	P_{w_1}	P_{w_2}	P_{w_3}
w_1	w_2	w_3	m_2	m_3	m_1
w_2	w_3	w_1	m_3	m_1	m_2
w_3	w_1	w_2	m_1	m_2	m_3
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

Table 4.1: Preference profile for Example 4.1

The outcome of the MPDA algorithm at P_N is

$$\mu_M = [(m_1, w_1), (m_2, w_2), (m_3, w_3)],$$

and that of the WPDA algorithm is

$$\mu_W = [(m_1, w_3), (m_2, w_1), (m_3, w_2)].$$

However, the outcome of the MWPDA algorithm with $\kappa = (2, 2, 2)$ is

$$\mu = [(m_1, w_2), (m_2, w_3), (m_3, w_1)].$$

Note that in μ_M, each man gets his best choice whereas each woman gets her worst, and conversely, in μ_W, each woman gets her best choice whereas each man gets his worst. However, in μ, all men and women get their second-best choices.

In view of this example, we define the notion of *minimum regret under stability*. This notion is introduced in Knuth (1976) as a desirable property of a matching.

Definition 4.1. Let P_N be a preference profile and let μ be a matching at P_N. Then, the ***regret*** of μ at P_N is defined as $\alpha(\mu, P_N) = \max\limits_{a \in N} \; rank(P_a, \mu(a))$.

The ***minimum regret under stability*** at P_N is defined as $\alpha(P_N) = \min\limits_{\mu \in \mathcal{C}(P_N)} \alpha(\mu, P_N)$.

It is worth mentioning that the regret of an unstable matching can be strictly less than the minimum regret under stability.

Definition 4.2. (Knuth, 1976) A matching μ^* is ***minimum regret stable*** at a preference profile P_N if it is stable at P_N and its regret is same as minimum regret under stability at P_N, that is, $\alpha(\mu^*, P_N) = \alpha(P_N)$.

An algorithm is called ***minimum regret stable*** if it produces a minimum regret stable matching at every preference profile.

It is worth noting that the minimum regret property has a close resemblance with a Rawlsian welfare function. Roughly speaking, this property tries to improve the outcome of the 'poorest of the poor' agent.

Clearly, both WPDA and MPDA algorithms do not satisfy this property in general since these algorithms always maximize the matches of one side of the market (women or men), and consequently maximizes the regret of the other side. For instance, consider Example 4.1. The regret of the both outcomes of the WPDA and MPDA algorithms is 3. However, the same of the outcome of the MWPDA algorithm with $\kappa = (2, 2, 2)$ is 2.

4.1 Sequential MWPDA algorithm

In this subsection, we present an algorithm that is minimum regret stable. We call this the ***sequential MWPDA*** algorithm. It involves a sequence of rounds. At every round, it performs an MWPDA algorithm with a cut-off vector. Below, we present a formal description of this algorithm at a preference profile P_N. Let $\kappa^* = \max\limits_{m \in M} \; rank(P_m, \mu_M(m))$.

Round 1. Perform the MWPDA algorithm with κ such that $\kappa_m = \kappa^*$ for all $m \in M$. Let μ_1^* be the outcome of the MWPDA algorithm at Round 1. If $rank(P_m, \emptyset) \leq \kappa^*$ for all $m \in M$ or $rank(P_w, \mu_1^*(w)) \leq \kappa^*$ for all $w \in W$, then conclude that the algorithm converges and define μ_1^* as the outcome of the sequential MWPDA algorithm. Else, go to Round 2.

Round 2. Perform the MWPDA algorithm with κ such that $\kappa_m = \kappa^* + 1$ for all $m \in M$. Let μ_2^* be the outcome of the MWPDA algorithm at Round 2. If $rank(P_m, \emptyset) \leq \kappa^* + 1$ for all $m \in M$ or $rank(P_w, \mu_2^*(w)) \leq \kappa^* + 1$ for all $w \in W$, then conclude that the algorithm converges and define μ_2^* as the outcome of the sequential MWPDA algorithm. Else, go to Round 3.

$$\vdots$$

Continue this till a round k such that either we have $rank(P_m, \emptyset) \leq \kappa^* + k - 1$ for all $m \in M$ or $rank(P_w, \mu_k^*(w)) \leq \kappa^* + k - 1$ for all $w \in W$ for the *first time* at Round k.[17] In other words, k is such that for all round $l < k$, there exists $m \in M$ with $rank(P_m, \emptyset) > \kappa^* + l - 1$ and $w \in W$ with $rank(P_w, \mu_l^*(w)) > \kappa^* + l - 1$. Define μ_k^* as the outcome of the sequential MWPDA algorithm.

Remark 4.1. It is worth noting that in order to execute the sequential MWPDA algorithm at a preference profile P_N, first one needs to compute the men-optimal stable matching at P_N.

Remark 4.2. By Theorem 3.3, the MWPDA algorithm used at every round of the sequential MWPDA algorithm converges at Stage 1. This ensures quick convergence of the sequential MWPDA algorithm.

Our next result says that the sequential MWPDA algorithm produces the women-optimal matching in the set of all minimum regret stable matchings.

[17] Since κ_m cannot be bigger than $q + 1$, such a round must exist.

Theorem 4.1. *The sequential MWPDA algorithm is minimum regret stable. Furthermore, the outcome of the sequential MWPDA algorithm is women-optimal in the set of all minimum regret stable matchings.*

The proof of this theorem is relegated to Appendix C.

5 Stable matching with forced and forbidden pairs

The notion of stable matching with *forced pairs* is introduced in Knuth (1976), and that with *forbidden pairs* is introduced in Dias et al. (2003). In this section, we provide an algorithm that produces stable matching with forced and forbidden pairs, whenever such a matching exists.

Definition 5.1. Given a set of pairs $Q_1 \subseteq M \times W$, we say a matching μ is with ***forced pairs*** Q_1 if every pair in Q_1 is matched in μ, that is, $\mu(m) = w$ for all $(m, w) \in Q_1$.

Definition 5.2. Given a set of pairs $Q_2 \subseteq M \times W$, we say a matching μ is with ***forbidden pairs*** Q_2 if no pair in Q_2 is matched in μ, that is, $\mu(m) \neq w$ for all $(m, w) \in Q_2$.

5.1 Conditional MWPDA algorithm

Consider a preference profile P_N and let Q_1 be a set of forced pairs and Q_2 be a set of forbidden pairs. Note that for all $(m, w), (m', w') \in Q_1$ with $(m, w) \neq (m', w')$, we have $m \neq m'$ and $w \neq w'$.[18] For $m \in M$, with slight abuse of notation, we say $m \in Q_1$, if there exists $w \in W$ such that $(m, w) \in Q_1$.

In what follows, we present an algorithm, called ***conditional MWPDA algorithm given*** (Q_1, Q_2), that produces a stable matching with forced pairs Q_1 and forbidden pairs Q_2, whenever such a matching exists. The algorithm involves a sequence of rounds. At every round, an MWPDA algorithm is performed with a cut-vector κ such that $\kappa_m = rank(P_m, w)$ for all $m \in Q_1$ with $(m, w) \in Q_1$. The cut-off parameters for other men may change over rounds; they are defined at the beginning of each round of the conditional MWPDA algorithm.

Round 1. Define κ^1 such that for all $m \notin Q_1$, $\kappa^1_m = rank(P_m, \emptyset)$. Perform the MWPDA algorithm with κ^1. Let μ^*_1 be the outcome of the MWPDA algorithm at Round 1.

(i) If μ^*_1 is with forced pairs Q_1 and forbidden pairs Q_2, then conclude that the algorithm converges and define μ^*_1 as the outcome of the algorithm.

(ii) Else, if there exists a pair $(m, w) \in Q_1$ such that $\mu^*_1(m) \neq w$, then conclude that the algorithm STOPS.

(iii) Else, go to Round 2.

[18]Otherwise there will be no stable matching with forced pairs Q_1.

Round 2. Define κ^2 such that for all $m \notin Q_1$,

$$\kappa_m^2 = \begin{cases} rank(P_m, \mu_1^*(m)) & \text{if } (m, \mu_1^*(m)) \notin Q_2; \\ rank(P_m, \mu_1^*(m)) - 1 & \text{if } (m, \mu_1^*(m)) \in Q_2. \end{cases}$$

Perform the MWPDA algorithm with κ^2. Let μ_2^* be the outcome of the MWPDA algorithm at Round 2.

(i) If μ_2^* is with forced pairs Q_1 and forbidden pairs Q_2, then conclude that the algorithm converges and define μ_2^* as the outcome of the algorithm.

(ii) Else, if there exists a pair $(m, w) \in Q_1$ such that $\mu_2^*(m) \neq w$ or if there exists $m \in M$ such that $rank(P_m, \mu_2^*(m)) > \kappa_m^2$, then conclude that the algorithm STOPS.

(iii) Else, go to Round 3.

$$\vdots$$

Note that for any two consecutive rounds r and $r+1$, for each $m \notin Q_1$, we have $\kappa_m^r \leq \kappa_m^{r+1}$, and for at least one $m \notin Q_1$, we have $\kappa_m^r < \kappa_m^{r+1}$. Therefore, if the algorithm does not converge or STOP at any round, then there will come a round r where some $m \notin Q_1$ will have $\kappa_m^r = 0$. In that case too, conclude that the algorithm STOPS.

5.2 Conditional MWPDA algorithm produces stable matching with forced and forbidden pairs

The following result says that a stable matching with given forced and forbidden pairs exists at a preference profile only if the conditional MWPDA algorithm converges at that preference profile. It further says that whenever the conditional MWPDA algorithm converges, it produces a stable matching with given forced and forbidden pairs, which is also women-optimal in the set of all stable matchings with the given forced and forbidden pairs. Thus, if at a preference profile, the conditional MWPDA algorithm STOPS at any round, then it must be that there is no stable matching with the corresponding forced and forbidden pairs at that preference profile.

Theorem 5.1. *A stable matching with forced pairs Q_1 and forbidden pairs Q_2 exists at a preference profile P_N if and only if the conditional MWPDA algorithm given (Q_1, Q_2) converges at P_N. Further, whenever this algorithm converges, the outcome is women-optimal in the set of all stable matchings with forced pairs Q_1 and forbidden pairs Q_2.*

The proof of this theorem is relegated to Appendix D.

By the construction of the conditional MWPDA algorithm, we obtain the following corollary from Theorem 5.1. It says that whenever there is no forbidden pair, the conditional MWPDA algorithm will come to a conclusion at the first round itself: either it will converge or it will STOP. If it converges at this round, then a stable matching with given forced pairs is produced as the outcome which is also women-optimal in the set of all such stable matchings. If it STOPS, then that means there is no such a stable matching.

Corollary 5.1. *Let P_N be a preference profile and let Q_1 be a set of forced pairs.*

(i) If there exists a stable matching with forced pairs Q_1 at P_N, then the conditional MWPDA algorithm given $(Q_1, \emptyset)$ at P_N converges at Round 1. Furthermore, the outcome is women-optimal in the set of all stable matchings with forced pairs Q_1.

(ii) If there is no stable matching with forced pairs Q_1 at P_N, then the conditional MWPDA algorithm given $(Q_1, \emptyset)$ at P_N STOPS at Round 1.

Appendix A Proof of Theorem 3.1

In all our proofs, for a given MWPDA algorithm at a preference profile P_N, we use the notation μ_k^s to denote the outcome obtained at Step k of the WPDA algorithm at Stage s of the given MWPDA algorithm, and the notation t^* to denote the last stage of the MWPDA algorithm. We make two observations which we will use in our proofs.

Observation A.1. *Consider a stage, say s, and two steps l and k with $l \leq k$ of the WPDA algorithm at Stage s of an MWPDA algorithm at a preference profile P_N. Then, it follows from the property of the WPDA algorithm that for all $m \in M$, we have $\mu_k^s(m) R_m \mu_l^s(m)$.*

Observation A.2. *Consider a stage, say s, of an MWPDA algorithm at a preference profile P_N. It follows from the property of the WPDA algorithm that μ^s is stable at the preference profile $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^s(w_1)}, \ldots, P_{w_q}^{M^s(w_q)})$.*[19]

Fix a preference profile P_N. Take an arbitrary cut-off vector κ and consider the MWPDA algorithm with κ at P_N. First, we prove a lemma that says that the match of a woman gets better over stages.

Lemma A.1. *For all $r \leq s \leq t^*$ and all $w \in W$, $\mu^s(w) R_w \mu^r(w)$.*

Proof of Lemma A.1. By the definition of the MWPDA algorithm, we have $\mu^s(w) R_w^{M^s(w)} \emptyset$ for all $w \in W$. This, together with the construction of $M^s(w)$, implies that $\mu^s(w) R_w \emptyset$ for all $w \in W$. So, if $\mu^r(w) = \emptyset$ for some $w \in W$, then there is nothing to show for that w. Take $w \in W$ such that $\mu^r(w) = m \in M$ and take $r < t^*$. It is enough to show that $\mu^{r+1}(w) R_w \mu^r(w)$. Assume for contradiction that $m P_w \mu^{r+1}(w)$.

[19]See Subsection 3.2 for the definition of the notation $P_w^{M^s(w)}$.

Because $\mu^r(m) = w$, by the definition of the MWPDA algorithm, we have $W^r(m) = W^{r+1}(m)$ and $w \in W^r(m)$. Combining all these, we have $w \in W^{r+1}(m)$, which implies $m \in M^{r+1}(w)$. Since $mP_w\mu^{r+1}(w)$ and $m \in M^{r+1}(w)$, we have $mP_w^{M^{r+1}(w)}\mu^{r+1}(w)$. By the definition of the MWPDA algorithm, there must be some step l of the WPDA algorithm at Stage $r+1$ where m rejects w to be tentatively matched with some $w' \in W^{r+1}(m)$ whom he prefers to w. This means

$$w'P_m w, \text{ and} \tag{A.1a}$$

$$mP_{w'}^{M^{r+1}(w')}\emptyset. \tag{A.1b}$$

Moreover, since $w' \in W^{r+1}(m)$ and $W^r(m) = W^{r+1}(m)$, we have $w' \in W^r(m)$.

Assume that Step l of the WPDA algorithm at Stage $r+1$ has the property that there is no $\hat{w} \in W$ with $\mu^r(\hat{w}) \neq \emptyset$ and $\mu^r(\hat{w})P_{\hat{w}}\mu^{r+1}(\hat{w})$ such that man $\mu^r(\hat{w})$ rejects woman $\hat{w}$ at some step $l' < l$ of the WPDA algorithm at Stage $r+1$. This is without loss of generality because, if there is such woman $\hat{w}$, then we can take $w = \hat{w}$.

Suppose $mP_{w'}\mu^r(w')$. Because $w' \in W^r(m)$, we have $m \in M^r(w')$. Since $mP_{w'}\mu^r(w')$ and $m \in M^r(w')$, it follows from the construction of $P_{w'}^{M^r(w')}$ that $mP_{w'}^{M^r(w')}\mu^r(w')$. This, together with (A.1a) and the fact $\mu^r(m) = w$, implies that (m, w') blocks μ^r at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^r(w_1)}, \ldots, P_{w_q}^{M^r(w_q)})$, which is a contradiction to Observation A.2. So, it must be that $\mu^r(w')R_{w'}m$. Because $\mu^r(w) = m$, $w \neq w'$, and $\mu^r(w')R_{w'}m$, we have $\mu^r(w')P_{w'}m$. Moreover, it follows from (A.1b) and the construction of $P_{w'}^{M^{r+1}(w')}$ that $mP_{w'}\emptyset$. Combining the facts that $\mu^r(w')P_{w'}m$ and $mP_{w'}\emptyset$, we have

$$\mu^r(w')P_{w'}mP_{w'}\emptyset. \tag{A.2}$$

Now, we complete the proof of the lemma. Because $w' \in W^{r+1}(m)$, we have $m \in M^{r+1}(w')$. Furthermore, (A.2) implies $\mu^r(w') \in M$. This, together with the definition of the MWPDA algorithm, yields $\mu^r(w') \in M^{r+1}(w')$. Since $m, \mu^r(w') \in M^{r+1}(w')$, it follows from (A.2) that $\mu^r(w')P_{w'}^{M^{r+1}(w')}mP_{w'}^{M^{r+1}(w')}\emptyset$. This, together with the fact that woman w' is tentatively matched with man m at Step l of the WPDA algorithm at Stage $r+1$, implies that $\mu^r(w')$ rejects w' at some step $l' < l$ of the WPDA algorithm at Stage $r+1$. However, this contradicts our assumption on Step l of the WPDA algorithm at Stage $r+1$, which completes the proof of Lemma A.1. ∎

Completion of the proof of Theorem 3.1. In view of Remark 2.1, we show that the outcome of the MWPDA algorithm is pairwise stable. Note that by the definition of the MWPDA algorithm, its outcome is always individually rational. We show that no pair can block its outcome. Let μ be the outcome of the MWPDA algorithm. Assume for contradiction that a pair $(m, w) \in M \times W$ blocks μ at P_N.

Since μ is individually rational at P_N and (m, w) is a blocking pair of μ at P_N, we have $wP_m\mu(m)R_m\emptyset$ and $mP_w\mu(w)R_w\emptyset$. Because $wP_m\mu(m)$, there must be some stage, say r^*, at which m proposes w for the first time. If $\mu^{r^*}(w)R_w m$, then by Lemma A.1, we have $\mu(w)R_w m$, which contradicts the fact $mP_w\mu(w)R_w\emptyset$. So, assume $mP_w\mu^{r^*}(w)$. Since $w \in W^{r^*}(m)$ and $mP_w\mu^{r^*}(w)$, w proposes m and gets rejected at some step, say l, of the WPDA algorithm at Stage r^*. Since $wP_m\emptyset$, by Observation A.1, this means

$$\mu^{r^*}(m)P_m w P_m \emptyset. \tag{A.3}$$

If $r^* = t^*$, then (A.3) implies $\mu(m)P_m w$, which contradicts the fact $wP_m\mu(m)R_m\emptyset$. So, assume $r^* < t^*$. By (A.3), we have $\mu^{r^*}(m) \neq \emptyset$. Since $r^* < t^*$ and $\mu^{r^*}(m) \neq \emptyset$, m proposes the women in $W^{r^*}(m)$ at the beginning of Stage r^*+1. Then, using a similar argument as for the derivation of (A.3), we have $\mu^{r^*+1}(m)P_m w P_m\emptyset$. Continuing in this manner, it follows that $\mu(m)P_m w P_m\emptyset$, which contradicts the fact $wP_m\mu(m)R_m\emptyset$. This completes the proof of Theorem 3.1. ■

Appendix B Proofs of Theorem 3.2 and Theorem 3.3

In this section, we prove Theorem 3.2 and Theorem 3.3. We prove Theorem 3.3 first since we use that in the proof of Theorem 3.2.

B.1 Proof of Theorem 3.3

We prove Theorem 3.3 using the following lemmas. Our first lemma is taken from McVitie and Wilson (1970). It says that the set of unmatched men or women stays the same in all stable matchings.

Lemma B.1. *(McVitie and Wilson, 1970) Let P_N be a preference profile and let $\mu, \mu' \in \mathcal{C}(P_N)$. Then, for all $a \in N$, $\mu(a) = \emptyset$ implies $\mu'(a) = \emptyset$.*

Our next lemma provides a sufficient condition on κ such that a given stable matching at a preference profile P_N remains stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$.

Lemma B.2. *Let P_N be a preference profile and let $\mu \in \mathcal{C}(P_N)$. Then, μ is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots,$ $P_{w_q}^{M^1(w_q)})$ if $\kappa_m \geq \min\left\{rank(P_m, \mu(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$.*

Proof of Lemma B.2. Suppose $\kappa_m \geq \min\left\{rank(P_m, \mu(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. In view of Remark 2.1, we show that μ is pairwise stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. First note that since $\kappa_m \geq \min\left\{rank(P_m, \mu(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$, we have $\mu(w) \in M^1(w) \cup \{\emptyset\}$ for all $w \in W$. Moreover, since $\mu(w) \in M^1(w) \cup \{\emptyset\}$ for all $w \in W$, we have for all $w \in W$ and all $m \in M$, $mR_w^{M^1(w)}\mu(w)$ implies $mR_w\mu(w)$. Further note that the preferences of the men are unchanged from P_N to $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Therefore, if (m, w) blocks μ at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$, then they also block μ at P_N contradicting the fact that μ is stable at P_N. Hence, μ cannot have a blocking pair at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Using a similar logic, it follows that μ is individually rational at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$.

■

Completion of the proof of Theorem 3.3. ***(If part)*** Take a cut-off vector κ such that $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. We show the MWPDA algorithm with κ at P_N converges at Stage 1. By the definition of the algorithm, it converges at Stage 1 if $W^1(m) = \mathcal{A}(P_m)$ for all $m \in M$ with $\mu^1(m) = \emptyset$. Take $m \in M$. If $\mu_M(m) = \emptyset$, then by the definition of κ, m proposes all acceptable women at the beginning of Stage 1, and hence $W^1(m) = \mathcal{A}(P_m)$. Suppose $\mu_M(m) \neq \emptyset$. It is enough to show that $\mu^1(m) \neq \emptyset$. Because $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$, by Lemma B.2, μ_M is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Furthermore, by Observation A.2, μ^1 is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Since μ^1 and μ_M both are stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$, by Lemma B.1, we have $\mu^1(m) \neq \emptyset$.

(Only-if part) Take a cut-off vector κ such that $\kappa_m < \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for some $m \in M$. Assume for contradiction that the MWPDA algorithm with κ at P_N converges at Stage 1. Since $\kappa_m < \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$, this means $\mu^1(m) \neq \emptyset$ and $rank(P_m, \mu^1(m)) \leq \kappa_m$. Combining the facts $rank(P_m, \mu^1(m)) \leq \kappa_m$ and $\kappa_m < \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$, we have $rank(P_m, \mu^1(m)) < rank(P_m, \mu_M(m))$. This, along with Remark 3.2, implies μ^1 is not stable at P_N, which contradicts Theorem 3.1. This completes the proof of Theorem 3.3. ■

B.2 Proof of Theorem 3.2

Proof of Theorem 3.2. Let μ_M be the men-optimal stable matching at P_N. Because $\mu \in \mathcal{C}(P_N)$, by Remark 3.2, we have $rank(P_m, \mu(m)) \geq rank(P_m, \mu_M(m))$ for all $m \in M$. This, together with the fact that $\kappa_m = rank(P_m, \mu(m))$ for all $m \in M$, means $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. Therefore, by Theorem 3.3, the MWPDA algorithm with κ converges at Stage 1.

Now, we show $\mu^1 = \mu$. Since $\kappa_m = rank(P_m, \mu(m))$ for all $m \in M$, we have $\kappa_m \geq \min\left\{rank(P_m, \mu(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. This, together with Lemma B.2, implies that μ is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Moreover, by the definition of the MWPDA algorithm, μ^1 is women-optimal stable matching at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Since $\mu \in \mathcal{C}(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$ and μ^1 is women-optimal stable matching at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$, by Remark 3.2, it follows that

$$\mu(m) R_m \mu^1(m) \text{ for all } m \in M. \tag{B.1}$$

Since $\mu, \mu^1 \in \mathcal{C}(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$, by Lemma B.1, we have

$$\mu^1(m) = \mu(m) \text{ for all } m \in M \text{ with } \mu^1(m) = \emptyset. \tag{B.2}$$

By the definition of the MWPDA algorithm, $rank(P_m, \mu^1(m)) \leq \kappa_m$ for all $m \in M$ with $\mu^1(m) \neq \emptyset$. This, together with definition of κ and (B.1), implies that

$$\mu^1(m) = \mu(m) \text{ for all } m \in M \text{ with } \mu^1(m) \neq \emptyset. \tag{B.3}$$

(B.2) and (B.3) together imply $\mu^1 = \mu$.

It remains to show that the MWPDA algorithm with κ converges at the first step of the WPDA algorithm at Stage 1. Suppose not. Then, there exists a pair (m, w) such that at the first step of the WPDA algorithm at Stage 1, w proposes m and gets rejected. By the definition of the MWPDA algorithm, this means $w \in W^1(m)$ and $mP_w^{M^1(w)}\mu^1(w)$. Moreover, since $\mu^1 = \mu$ and $mP_w^{M^1(w)}\mu^1(w)$, we have $\mu(m) \neq w$. The facts $\kappa_m = rank(P_m, \mu(m))$, $w \in W^1(m)$, and $w \neq \mu(m)$ together imply $wP_m\mu(m)$. Because $\mu^1 = \mu$, this, together with the fact $mP_w^{M^1(w)}\mu^1(w)$, implies (m, w) blocks μ^1 at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$, a contradiction to Observation A.2. This completes the proof of Theorem 3.2. ■

Appendix C Proof of Theorem 4.1

We prove a sequence of lemmas that we use in the proof of Theorem 4.1.

Lemma C.1. *Let P_N be a preference profile and let κ be such that $\kappa_m \geq rank(P_m, \mu_M(m))$ for all $m \in M$. Suppose μ is the outcome of the MWPDA algorithm with κ at P_N. Then, $rank(P_m, \mu(m)) \leq \kappa_m$ for all $m \in M$.*

Proof of Lemma C.1. By Theorem 3.1, $\mu \in \mathcal{C}(P_N)$. Since $\mu, \mu_M \in \mathcal{C}(P_N)$, by Lemma B.1, we have $\mu(m) = \mu_M(m)$ for all $m \in M$ with $\mu(m) = \emptyset$. This, together with the definition of κ, implies

$$rank(P_m, \mu(m)) \leq \kappa_m \text{ for all } m \in M \text{ with } \mu(m) = \emptyset. \tag{C.1}$$

By the definition of κ, we have $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. Therefore, by Theorem 3.3, the MWPDA algorithm with κ at P_N converges at Stage 1 producing μ. This, together with the definition of the MWPDA algorithm, implies

$$rank(P_m, \mu(m)) \leq \kappa_m \text{ for all } m \in M \text{ with } \mu(m) \neq \emptyset. \tag{C.2}$$

The proof of Lemma C.1 follows from (C.1) and (C.2). ■

The implication of our next lemma is as follows. Let μ be the outcome of the MWPDA algorithm with cut-off vector κ where κ is such that every man gets to propose the woman (together with other women) who he would be matched with in the men-optimal stable matching (if a man is unmatched in the men-optimal stable matching, then he proposes all acceptable women). Let μ' be another stable matching where the rank of the match of every man m (the match might be some woman or $\emptyset$) according to P_m is less than or equal to κ_m. Then, for every woman, the match in μ must be at least as good as that in μ'.

Lemma C.2. *Let P_N be a preference profile and let κ be such that $\kappa_m \geq rank(P_m, \mu_M(m))$ for all $m \in M$. Let μ be the outcome of the MWPDA algorithm with κ at P_N. Suppose $\mu' \in \mathcal{C}(P_N)$ is such that $rank(P_m, \mu'(m)) \leq \kappa_m$ for all $m \in M$. Then, $\mu(w) R_w \mu'(w)$ for all $w \in W$.*

Proof of Lemma C.2. Suppose μ and μ' are as defined in Lemma C.2. Since $\kappa_m \geq rank(P_m, \mu_M(m))$ for all $m \in M$, we have $\kappa_m \geq \min\left\{rank(P_m, \mu_M(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. This, along with Theorem 3.3, implies that the MWPDA algorithm with κ at P_N converges at Stage 1 producing μ. By Observation A.2, this means μ is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Also, since $rank(P_m, \mu'(m)) \leq \kappa_m$ for all $m \in M$, we have $\kappa_m \geq \min\left\{rank(P_m, \mu'(m)), \max\left\{|\mathcal{A}(P_m)|, 1\right\}\right\}$ for all $m \in M$. This, along with Lemma B.2, implies that μ' is stable at $(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$. Because μ, $\mu' \in \mathcal{C}(P_{m_1}, \ldots, P_{m_p}, P_{w_1}^{M^1(w_1)}, \ldots, P_{w_q}^{M^1(w_q)})$ and μ is the outcome of the WPDA algorithm at Stage 1 of the MWPDA algorithm, by Remark 3.2, we have $\mu(w) R_w^{M^1(w)} \mu'(w)$ for all $w \in W$. By the definition of the MWPDA algorithm, $\mu(w) \in M^1(w) \cup \{\emptyset\}$. As $rank(P_m, \mu'(m)) \leq \kappa_m$ for all $m \in M$, we have $\mu'(w) \in M^1(w) \cup \{\emptyset\}$ for all $w \in W$. Since for all $w \in W$, we have $\mu(w), \mu'(w) \in M^1(w) \cup \{\emptyset\}$ and $\mu(w) R_w^{M^1(w)} \mu'(w)$, by the construction of $P_w^{M^1(w)}$, we have $\mu(w) R_w \mu'(w)$ for all $w \in W$. This completes the proof of Lemma C.2. $\blacksquare$

Completion of the proof of Theorem 4.1. By Theorem 3.1, it is straightforward that the sequential MWPDA algorithm is stable. We proceed to show that the sequential MWPDA algorithm produces a minimum regret stable matching at every preference profile. Take a preference profile P_N. Let κ be the cut-off vector that is used at the terminal round of the sequential MWPDA algorithm at P_N and μ be the outcome of the sequential MWPDA algorithm at P_N. It follows from the definition of the sequential MWPDA algorithm that $\kappa_m \geq rank(P_m, \mu_M(m))$ for all $m \in M$. Therefore, by Lemma C.1 along with the definition of the sequential MWPDA algorithm, we have

$$rank(P_m, \mu(m)) \leq \kappa_m \text{ for all } m \in M. \tag{C.3}$$

Claim C.1. *$\kappa_m \leq \alpha(P_N)$ for all $m \in M$.*

Proof of Claim C.1. Assume for contradiction that $\kappa_m > \alpha(P_N)$ for some (and hence, all) $m \in M$. Consider the round of the sequential MWPDA algorithm where the MWPDA algorithm is performed with $\hat{\kappa}$ where $\hat{\kappa}_m = \alpha(P_N)$ for all $m \in M$. Let $\hat{\mu}$ be the outcome of that round. By the definition of $\alpha(P_N)$, there must exist $\mu' \in \mathcal{C}(P_N)$ such that $\alpha(\mu', P_N) = \alpha(P_N)$. Because $\alpha(\mu', P_N) = \alpha(P_N)$, we have $rank(P_m, \mu'(m)) \leq \alpha(P_N)$ for all $m \in M$. By Lemma C.2, this means $\hat{\mu}(w) R_w \mu'(w)$ for all $w \in W$. Therefore, $\max\limits_{w \in W}\; rank(P_w, \hat{\mu}(w)) \leq \max\limits_{w \in W}\; rank(P_w, \mu'(w)) \leq \alpha(P_N)$. By the definition of the sequential MWPDA algorithm, this means that the algorithm cannot go for another round, which contradicts the fact that $\kappa_m > \alpha(P_N)$ for all $m \in M$. This completes the proof of Claim C.1. $\square$

Since κ is the cut-off vector that is used at the terminal round of the sequential MWPDA algorithm at P_N and μ is the outcome of the sequential MWPDA algorithm at P_N, one of the following two statements must hold.

(1) $rank(P_m, \emptyset) \leq \kappa_m$ for all $m \in M$.

(2) $rank(P_w, \mu(w)) \leq \kappa_m$ for all $w \in W$ and for some (and hence, all) $m \in M$.

We distinguish the following two cases.

CASE 1: Suppose $rank(P_m, \emptyset) \leq \kappa_m$ for all $m \in M$.
Since $rank(P_m, \emptyset) \leq \kappa_m$ for all $m \in M$ and μ is the outcome of the sequential MWPDA, it is easy to verify that μ is the women-optimal stable matching at P_N. By the definition of $\alpha(P_N)$, there must exist $\mu' \in \mathcal{C}(P_N)$ such that $\alpha(\mu', P_N) = \alpha(P_N)$. Since μ is the women-optimal stable matching, we have $rank(P_w, \mu(w)) \leq rank(P_w, \mu'(w)) \leq \alpha(P_N)$ for all $w \in W$. Moreover, by Claim C.1 along with (C.3), we have $rank(P_m, \mu(m)) \leq \alpha(P_N)$ for all $m \in M$. Combining the facts that $rank(P_m, \mu(m)) \leq \alpha(P_N)$ for all $m \in M$ and $rank(P_w, \mu(w)) \leq \alpha(P_N)$ for all $w \in W$, we have $\alpha(\mu, P_N) \leq \alpha(P_N)$. By the definition of $\alpha(P_N)$, this means $\alpha(\mu, P_N) = \alpha(P_N)$. So, μ is a minimum regret stable matching at P_N. Because μ is the women-optimal stable matching at P_N, this implies that μ is women-optimal in the set of all minimum regret stable matchings at P_N.

CASE 2: Suppose $rank(P_w, \mu(w)) \leq \kappa_m$ for all $w \in W$ and for some (and hence, all) $m \in M$.
Since $rank(P_w, \mu(w)) \leq \kappa_m$ for all $w \in W$ and for some $m \in M$, it follows from (C.3) and the definition of the sequential MWPDA algorithm that $\alpha(\mu, P_N) \leq \kappa_m$ for all $m \in M$. This, together with Claim C.1, implies that $\alpha(\mu, P_N) \leq \kappa_m \leq \alpha(P_N)$ for all $m \in M$. By the definition of $\alpha(P_N)$, this means

$$\alpha(\mu, P_N) = \kappa_m = \alpha(P_N) \text{ for all } m \in M. \tag{C.4}$$

By (C.4), we have $\alpha(\mu, P_N) = \alpha(P_N)$. So, μ is a minimum regret stable matching at P_N.

Let μ' be a minimum regret stable matching at P_N. Clearly, $rank(P_m, \mu'(m)) \leq \alpha(P_N)$ for all $m \in M$. This, together with (C.4), implies that $rank(P_m, \mu'(m)) \leq \kappa_m$ for all $m \in M$. Furthermore, it follows from the definition of the sequential MWPDA algorithm that μ is the outcome of the MWPDA algorithm with κ at P_N. Since $\kappa_m \geq rank(P_m, \mu_M(m))$ for all $m \in M$, μ is the outcome of the MWPDA algorithm with κ, and μ' is a stable matching with $rank(P_m, \mu'(m)) \leq \kappa_m$ for all $m \in M$, by Lemma C.2, we have $\mu(w) R_w \mu'(w)$ for all $w \in W$. Since μ is a minimum regret stable matching at P_N, this implies that μ is women-optimal in the set of all minimum regret stable matchings at P_N.

Since Case 1 and Case 2 are exhaustive, it follows that the outcome of the sequential MWPDA algorithm is women-optimal in the set of all minimum regret stable matchings. This completes the proof of Theorem 4.1. ■

Appendix D Proof of Theorem 5.1

The following lemma follows from Lemma 1 in Gale and Sotomayor (1985), which establishes a relationship between two stable matchings at a preference profile.

Lemma D.1. *Let P_N be a preference profile and let $\mu, \mu' \in \mathcal{C}(P_N)$. Then, $\mu(m) R_m \mu'(m)$ for all $m \in M$ if and only if $\mu'(w) R_w \mu(w)$ for all $w \in W$.*

Let us first recall some of the notations used in the context of the conditional MWPDA algorithm. For a preference profile P_N, a set of forced pairs Q_1, and a set of forbidden pairs Q_2, κ^r is the cut-off vector associated with the MWPDA algorithm at Round r of the conditional MWPDA algorithm given (Q_1, Q_2) and μ_r^* is the outcome of the MWPDA algorithm at Round r.

Completion of the proof of Theorem 5.1. It is obvious that if the conditional MWPDA algorithm given (Q_1, Q_2) converges at P_N, then there exists a stable matching with forced pairs Q_1 and forbidden pairs Q_2. We proceed to prove the rest of the theorem. Suppose there exists a stable matching with forced pairs Q_1 and forbidden pairs Q_2 at P_N. Let $\bar{\mathcal{C}}(P_N)$ be the set of all stable matchings at P_N with forced pairs Q_1 and forbidden pairs Q_2. Clearly, $\bar{\mathcal{C}}(P_N) \neq \emptyset$. Define the mapping $\mu^* : N \to N \cup \{\emptyset\}$ such that

(i) for all $m \in M$, $\mu^*(m) = x$ if and only if there exists a $\mu \in \bar{\mathcal{C}}(P_N)$ such that $\mu(m) = x$ and $\mu'(m) R_m x$ for all $\mu' \in \bar{\mathcal{C}}(P_N)$, and

(ii) for all $w \in W$, $\mu^*(w) = y$ if and only if there exists a $\mu \in \bar{\mathcal{C}}(P_N)$ such that $\mu(w) = y$ and $y R_w \mu'(w)$ for all $\mu' \in \bar{\mathcal{C}}(P_N)$.

It follows from the construction of μ^* that it is women-optimal in $\bar{\mathcal{C}}(P_N)$ (see Knuth (1976) for details). We show that the conditional MWPDA algorithm given (Q_1, Q_2) converges at P_N producing μ^* as the outcome.

If $\mu_1^* = \mu^*$, then we are done. Suppose $\mu_1^* \neq \mu^*$.

Claim D.1. *For all $m \in M$, we have*

(i) $rank(P_m, \mu_1^*(m)) \leq \kappa_m^1$, and

(ii) $\mu^*(m) R_m \mu_1^*(m)$.

Proof of Claim D.1. By the definition of κ^1, we have $\kappa_m^1 \geq rank(P_m, \mu^*(m))$ for all $m \in M$. Since $\mu^* \in \mathcal{C}(P_N)$, by Remark 3.2, we have $rank(P_m, \mu^*(m)) \geq rank(P_m, \mu_M(m))$ for all $m \in M$. Combining the facts that $\kappa_m^1 \geq rank(P_m, \mu^*(m))$ for all $m \in M$ and $rank(P_m, \mu^*(m)) \geq rank(P_m, \mu_M(m))$ for all $m \in M$, we have $\kappa_m^1 \geq rank(P_m, \mu_M(m))$ for all $m \in M$. Therefore, by Lemma C.1, $rank(P_m, \mu_1^*(m)) \leq \kappa_m^1$ for all $m \in M$. This proves (i) in Claim D.1.

By Lemma C.2, $\kappa_m^1 \geq rank(P_m, \mu^*(m))$ for all $m \in M$ implies $\mu_1^*(w) R_w \mu^*(w)$ for all $w \in W$. By Lemma D.1, this implies $\mu^*(m) R_m \mu_1^*(m)$ for all $m \in M$. This proves (ii) in Claim D.1.

$\square$

Claim D.2. *$\mu_1^*(m) = \mu^*(m) = w$ for all $(m, w) \in Q_1$.*

Proof of Claim D.2. Since $\kappa_m^1 = rank(P_m, w)$ for all $(m, w) \in Q_1$, $\mu^*(m) = w$ for all $(m, w) \in Q_1$, by Claim D.1, we have $\mu_1^*(m) = w$ for all $(m, w) \in Q_1$, which completes the proof of Claim D.2. $\square$

By Claim D.2, it follows that the conditional MWPDA algorithm given (Q_1, Q_2) will not stop at Round 1, and because it does not converge either at Round 1, it will go to Round 2.

Claim D.3. *$\kappa_m^2 \geq rank(P_m, \mu^*(m))$ for all $m \in M$.*

Proof of Claim D.3. By the definition of κ^2, we have $\kappa_m^2 = rank(P_m, \mu^*(m))$ for all $m \in Q_1$. Take $m \notin Q_1$. If $(m, \mu_1^*(m)) \notin Q_2$, then by the definition of κ^2 and (ii) in Claim D.1, we have $\kappa_m^2 \geq rank(P_m, \mu^*(m))$. On the other hand, if $(m, \mu_1^*(m)) \in Q_2$, which in particular means $\mu^*(m) \neq \mu_1^*(m)$, then by (ii) in Claim D.1, it must be that $\mu^*(m) P_m \mu_1^*(m)$. Therefore, by the definition of κ^2 and (ii) in Claim D.1, we have $\kappa_m^2 \geq rank(P_m, \mu^*(m))$. This completes the proof of Claim D.3. $\square$

Using similar logic as for Claims D.1 and D.2, it follows that

$$rank(P_m, \mu_2^*(m)) \leq \kappa_m^2 \text{ for all } m \in M, \tag{D.1a}$$

$$\mu^*(m) R_m \mu_2^*(m) \text{ for all } m \in M, \text{ and} \tag{D.1b}$$

$$\mu_2^*(m) = \mu^*(m) = w \text{ for all } (m, w) \in Q_1. \tag{D.1c}$$

Claim D.4. *$\mu_2^*(m) R_m \mu_1^*(m)$ for all $m \in M$ and there exists $m' \notin Q_1$ such that $\mu_2^*(m') P_{m'} \mu_1^*(m')$.*

Proof of Claim D.4. By the definition of κ^2, (D.1a) implies $\mu_2^*(m) R_m \mu_1^*(m)$ for all $m \notin M$. Moreover, as $\mu_1^* \neq \mu^*$, there must exist $m' \notin Q_1$ such that $(m', \mu_1^*(m')) \in Q_2$. This, together with the definition of κ^2 and (D.1a), yields $\mu_2^*(m') P_{m'} \mu_1^*(m')$. $\square$

By Claim D.4, (D.1a), and (D.1c), it follows that the conditional MWPDA algorithm given (Q_1, Q_2) either converges at Round 2 or goes to Round 3. If it goes to Round 3, then using similar logic as for Claim D.2, we have $\mu_3^*(m) = \mu^*(m) = w$ for all $(m, w) \in Q_1$, and that for Claim D.4, we have $\mu_3^*(m) R_m \mu_2^*(m)$ for all $m \in M$ and there exists $\bar{m} \notin Q_1$ such that $\mu_3^*(\bar{m}) P_{\bar{m}} \mu_2^*(\bar{m})$.

We argue that the conditional MWPDA algorithm given (Q_1, Q_2) must converge at some round.[20] Suppose not. Then, we will get a sequence of stable matchings $\mu_1^*, \mu_2^*, \ldots$ such that $\mu^*(m) R_m \ldots R_m \mu_2^*(m) R_m \mu_1^*(m)$ for all $m \in M$. Because $\mu_1^*, \mu_2^*, \ldots$ are all distinct and the number of stable matchings is finite, it follows that there must be a round where μ^* will be produced, and hence the conditional MWPDA algorithm will converge.

[20] Recall that the conditional MWPDA algorithm always terminates, that is, either converges or STOPS at every preference profile (see Subsection 5.1 for details).

Now, we show that the outcome of the conditional MWPDA algorithm given (Q_1, Q_2) is always μ^*. Let $\tilde{r}$ be the terminal round of the conditional MWPDA algorithm given (Q_1, Q_2). Using similar logic as for Claim D.1, we have $\mu^*(m) R_m \mu^*_{\tilde{r}}(m)$ for all $m \in M$. Since $\mu^*, \mu^*_{\tilde{r}} \in \mathcal{C}(P_N)$, by Lemma D.1, $\mu^*_{\tilde{r}}(w) R_w \mu^*(w)$ for all $w \in W$. Moreover, since the conditional MWPDA algorithm converges, it must be that $\mu^*_{\tilde{r}} \in \bar{\mathcal{C}}(P_N)$. Since $\mu^*_{\tilde{r}} \in \bar{\mathcal{C}}(P_N)$ and $\mu^*_{\tilde{r}}(w) R_w \mu^*(w)$ for all $w \in W$, by the definition of μ^*, we have $\mu^* = \mu^*_{\tilde{r}}$. This completes the proof of Theorem 5.1. ■

References

[1] Atila Abdulkadiroglu and Tayfun Sönmez. Matching markets: Theory and practice. *Advances in Economics and Econometrics*, 1:3–47, 2013.

[2] Vânia MF Dias, Guilherme D Da Fonseca, Celina MH De Figueiredo, and Jayme L Szwarcfiter. The stable marriage problem with restricted pairs. *Theoretical Computer Science*, 306(1-3):391–405, 2003.

[3] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[4] David Gale and Marilda Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11(3):223–232, 1985.

[5] Dan Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, 16(1):111–128, 1987.

[6] Dan Gusfield and Robert W Irving. *The stable marriage problem: structure and algorithms*. MIT press, 1989.

[7] Robert W Irving and Paul Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, 15(3):655–667, 1986.

[8] Alexander S Kelso Jr and Vincent P Crawford. Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, pages 1483–1504, 1982.

[9] Donald E Knuth. Mariages stables et leurs relations avec d'autres problèmes combinatoires. English translation in "Stable marriage and its relation to other combinatorial problems", Volume 10 of CRM Proceedings and Lecture Notes, American Mathematical Society, 1997. *Les Presses de l'Université de Montréal*, 1976.

[10] Ruth Martınez, Jordi Massó, Alejandro Neme, and Jorge Oviedo. An algorithm to compute the full set of many-to-many stable matchings. *Mathematical Social Sciences*, 47(2):187–210, 2004.

[11] David G McVitie and Leslie B Wilson. The stable marriage problem. *Communications of the ACM*, 14 (7):486–490, 1971.

[12] DG McVitie and Leslie B Wilson. Stable marriage assignment for unequal sets. *BIT Numerical Mathematics*, 10(3):295–309, 1970.

[13] Alvin E Roth. The college admissions problem is not equivalent to the marriage problem. *Journal of economic Theory*, 36(2):277–288, 1985.

[14] Alvin E Roth and Marilda Sotomayor. The college admissions problem revisited. *Econometrica: Journal of the Econometric Society*, pages 559–570, 1989.

[15] Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992.