Maximum Stable Matching with One-Sided Ties of Bounded Length

Chi-Kit Lam and C. Gregory Plaxton

University of Texas at Austin, Austin, TX 78712, USA
{geocklam, plaxton}@cs.utexas.edu

Abstract. We study the problem of finding maximum weakly stable matchings when preference lists are incomplete and contain one-sided ties of bounded length. We show that if the tie length is at most L, then it is possible to achieve an approximation ratio of $1 + (1 - \frac{1}{L})^L$. We also show that the same ratio is an upper bound on the integrality gap, which matches the known lower bound. In the case where the tie length is at most 2, our result implies an approximation ratio and integrality gap of $\frac{5}{4}$, which matches the known UG-hardness result.

Keywords: Stable matching · Approximation algorithm · Integrality gap.

1 Introduction

The stable matching model of Gale and Shapley [4] involves a two-sided market in which the agents are typically called men and women. Each agent has ordinal preferences over the agents on the opposite side. A matching is said to be stable if no man and woman prefer each other to their partners. Stable matchings always exist and can be computed efficiently by the proposal algorithm of Gale and Shapley. Their algorithm is also applicable when the preference lists are incomplete. In other words, agents are allowed to omit from their preference lists any unacceptable agent on the opposite side. If ties are allowed in the preference lists, the notion of stability can be generalized in several ways [9]. This paper focuses on weakly stable matchings, which always exist and can be obtained by invoking the Gale-Shapley algorithm after breaking all the ties arbitrarily. When incomplete lists are absent, every weakly stable matching is a maximum matching and hence has the same size. When ties are absent, the Rural Hospital Theorem guarantees that all stable matchings have the same size [5, 17]. However, when both ties and incomplete lists are present, weakly stable matchings can vary in size.

The problem of finding maximum weakly stable matchings with ties and incomplete lists has been studied in various settings. When ties and incomplete lists are allowed on both sides, there exist polynomial-time algorithms [11, 14, 15] that achieve an approximation ratio of $\frac{3}{2}$ $(= 1.5)$. Meanwhile, it is known [20] that getting an approximation ratio of $\frac{33}{29} - \varepsilon$ (≈ 1.1379) is NP-hard, and that getting an approximation ratio $\frac{4}{3} - \varepsilon$ (≈ 1.3333) is UG-hard. These hardness results hold in the case of two-sided ties even when the maximium tie length is two. The associated linear programming (LP) formulation has an integrality gap of at least $\frac{3L-2}{2L-1}$, where L is the maximum tie length [10].

For the case where ties appear only on one side of the market, algorithms with better approximation ratios have been developed using an LP-based approach [3, 10, 12] or the idea of rounding half-integral stable matchings [1, 8, 16]. The current best approximation ratio of $1 + \frac{1}{e}$ (≈ 1.3679) is attained by the LP-based algorithm that the authors recently presented in [12]. Meanwhile, it is known [7] that getting an approximation ratio of $\frac{21}{19} - \varepsilon$ (≈ 1.1053) is NP-hard, and that getting an approximation ratio of $\frac{5}{4} - \varepsilon$ (≈ 1.25) is UG-hard. These hardness results hold in the case of one-sided ties even when the maximum tie length is two. The associated LP formulation has an integrality gap of at least $1 + (1 - \frac{1}{L})^L$, where L is the maximum tie length [10]. Furthermore, for the case of one-sided ties with unbounded tie length, the integrality gap equals $1 + \frac{1}{e}$ and matches the attainable approximation ratio [12].

For the case of two-sided ties where the maximum tie length is two, Chiang and Pashkovich [2] showed that the algorithm of Huang and Kavitha [8] attains an approximation ratio of $\frac{4}{3}$ (≈ 1.3333), which matches the UG-hardness result [20] and the lower bound of the integrality gap [10]. A couple of results [6, 7] are known for the case of one-sided ties with bounded tie length, but they are subsumed by the approximation ratio of $1 + \frac{1}{e}$ for the case of one-sided ties with unbounded tie length.

Our Techniques and Contributions. In this paper, we focus on the problem of finding maximum weakly stable matchings with one-sided ties and incomplete lists when the tie length is bounded. We show that the algorithm of [12] achieves an approximation ratio of $1 + (1 - \frac{1}{L})^L$, where L is the maximum tie length. We also show that the same ratio is an upper bound on the integrality gap, which matches the lower bound of Iwama et al. [10]. For the case where $L = 2$, our result implies an approximation ratio and integrality gap of $\frac{5}{4}$, which matches the UG-hardness result of Halldórsson et al. [7].

Our analysis is based on four key properties established in [12]. Using these key properties, we extend the analysis of the approximation ratio to the case of bounded tie length. Moreover, we present a new, simpler charging argument. The main idea is to decompose the LP solution associated with each man-woman pair into a charge incurred by the man and a charge incurred by the woman based on an exchange function. We derive an upper bound for the charges incurred by a man using the strict ordering of his preferences, and an upper bound for the charges incurred by a woman using the bounded tie length assumption. By choosing a good exchange function, we show that every matched couple incurs a total charge of at most $1 + (1 - \frac{1}{L})^L$, providing an upper bound on the approximation ratio.

In Section 2, we review the key properties of the algorithm of [12] after presenting the stable matching model and its LP formulation. In Section 3, we present our simpler charging argument and use it to analyze the approximation ratio for the case of bounded tie length.

2 Stable Matching with One-Sided Ties

2.1 The Model

The formal definition of the stable matching problem with one-sided ties and incomplete lists (SMOTI) below follows the notations of [12].

In SMOTI, there are a set I of men and a set J of women. We assume that the sets I and J are disjoint and do not contain the element 0, which we use to denote being unmatched. Each man $i \in I$ has a preference relation $\geq_i$ over the set $J \cup \{0\}$ that satisfies antisymmetry, transitivity, and totality. Each woman $j \in J$ has a preference relation $\geq_j$ over the set $I \cup \{0\}$ that satisfies transitivity and totality. We denote this SMOTI instance as $(I, J, \{\geq_i\}_{i \in I}, \{\geq_j\}_{j \in J})$.

For every man $i \in I$ and woman $j \in J$, man i is said to be acceptable to woman j if $i \geq_j 0$. Similarly, woman j is said to be acceptable to man i if $j \geq_i 0$. The preference lists are allowed to be incomplete. In other words, there may exist $i \in I$ and $j \in J$ such that $0 >_j i$ or $0 >_i j$.

Notice that the preference relations $\{\geq_j\}_{j \in J}$ of the women are not required to be antisymmetric, while the preference relations $\{\geq_i\}_{i \in I}$ of the men are required to be antisymmetric. For every man $i \in I$, we write $>_i$ to denote the asymmetric part of $\geq_i$. For every woman $j \in J$, we write $>_j$ and $=_j$ to denote the asymmetric part and the symmetric part of $\geq_j$, respectively. A *tie* in the preference list of woman j is an equivalence class of size at least 2 with respect to the equivalence relation $=_j$, and the *length* of a tie is the size of this equivalence class.[1] We assume that there is at least one tie in the SMOTI instance, for otherwise every stable matching has the same size. We use L to denote the maximum length of the ties in the preference lists of the women, where $2 \leq L \leq |I| + 1$.

A matching is a subset $\mu \subseteq I \times J$ such that for every $(i, j), (i', j') \in \mu$, we have $i = i'$ if and only if $j = j'$. For every man $i \in I$, if $(i, j) \in \mu$ for some woman $j \in J$, we say that man i is matched to woman j in matching μ, and we write $\mu(i) = j$. Otherwise, we say that man i is unmatched in matching μ, and we write $\mu(i) = 0$. Similarly, for every woman $j \in J$, if $(i, j) \in \mu$ for some man $i \in I$, we say that woman j is matched to man i in matching μ, and we write $\mu(j) = i$. Otherwise, we say that woman j is unmatched in matching μ, and we write $\mu(j) = 0$.

A matching μ is *individually rational* if for every $(i, j) \in \mu$, we have $j \geq_i 0$ and $i \geq_j 0$. An individually rational matching μ is *weakly stable* if for every man $i \in I$ and woman $j \in J$, either $\mu(i) \geq_i j$ or $\mu(j) \geq_j i$. Otherwise, (i, j) forms a *strongly blocking pair*.

[1] Some of the literature on stable matching with indifferences does not allow an agent to be indifferent between being matched to an agent and being unmatched. Our formulation of the SMOTI problem allows for this possibility, since we can have $i =_j 0$ for any man i and woman j.

The goal of the maximum stable matching problem with one-sided ties and incomplete lists is to find a maximum-cardinality weakly stable matching for a given SMOTI instance.

2.2 The LP Formulation

The following LP formulation is based on that of Rothblum [18], which extends that of Vande Vate [19].

$$\begin{aligned}
\text{maximize} \quad & \sum_{(i,j)\in I\times J} x_{i,j} & & \\
\text{subject to} \quad & \sum_{j\in J} x_{i,j} \le 1 & & \forall i \in I && \text{(C1)}\\
& \sum_{i\in I} x_{i,j} \le 1 & & \forall j \in J && \text{(C2)}\\
& \sum_{\substack{j'\in J\\ j'>_i j}} x_{i,j'} + \sum_{\substack{i'\in I\\ i'\ge_j i}} x_{i',j} \ge 1 & & \forall (i,j)\in I\times J \text{ such that } j >_i 0 \text{ and } i >_j 0 && \text{(C3)}\\
& x_{i,j} = 0 & & \forall (i,j)\in I\times J \text{ such that } 0 >_i j \text{ or } 0 >_j i && \text{(C4)}\\
& x_{i,j} \ge 0 & & \forall (i,j)\in I\times J && \text{(C5)}
\end{aligned}$$

It is known [12, 18] that an integral solution $\mathbf{x} = \{x_{i,j}\}_{(i,j)\in I\times J}$ corresponds to the indicator variables of a weakly stable matching if and only if $\mathbf{x}$ satisfies constraints (C1)–(C5).

Given $\mathbf{x}$ which satisfies constraints (C1)–(C5), it is useful to define auxiliary variables

$$w_{i,j} = \begin{cases} 1 & \text{if } j = 0 \\ \sum_{\substack{j'\in J\\ j'>_i j}} x_{i,j'} & \text{if } j \neq 0 \end{cases}$$

for every $(i,j) \in I \times (J \cup \{0\})$, and

$$z_{i,j} = \sum_{\substack{i'\in I\\ i>_j i'}} x_{i',j}$$

for every $(i,j) \in (I \cup \{0\}) \times J$. The following lemma presents some simple properties of the auxiliary variables; see [13] for a proof.

Lemma 1. *The auxiliary variables satisfy the following conditions.*

(1) For every $i \in I$ and $j \in J$, we have $w_{i,j} + x_{i,j} \le 1$.
(2) For every $i \in I$ and $j, j' \in J$ such that $j >_i j'$, we have $w_{i,j} + x_{i,j} \le w_{i,j'}$.
(3) For every $i, i' \in I \cup \{0\}$ and $j \in J$ such that $i =_j i'$, we have $z_{i,j} = z_{i',j}$.
(4) For every $i \in I$ and $j \in J$ such that $j \ge_i 0$ and $i \ge_j 0$, we have $z_{i,j} \le w_{i,j}$.

2.3 The LP-Based Algorithm

Using the LP formulation of Section 2.2, the authors have previously established in [12] that there exists a polynomial-time algorithm with an approximation ratio of $1 + \frac{1}{e}$. The algorithm is based on a proposal process in which every man i maintains a priority p_i that gradually increases from 0 to 1. Between two successive increases of the priority of a man i, he attempts to propose to the set of women $\{j \in J : j \geq_i 0 \text{ and } p_i \geq w_{i,j}\}$ in decreasing order of his preference, where $w_{i,j}$ is the auxiliary variable corresponding to a fixed optimal fractional solution $\mathbf{x}$ of the LP. Each woman compares the men based on her preferences and breaks the ties by favoring men with higher priorities. The algorithm simulates this process in which the step size of the priority increases is infinitesimally small. More precisely, the algorithm runs in polynomial time and produces a weakly stable matching μ and priority values $\mathbf{p} = \{p_i\}_{i \in I}$ satisfying the following key properties [12, Lemmas 3.1 and 3.3].

(P1) Let $(i, j) \in \mu$. Then $j \geq_i 0$ and $i \geq_j 0$.
(P2) Let $i \in I$ be a man and $j \in J$ be a woman such that $j \geq_i \mu(i)$ and $i \geq_j 0$. Then $\mu(j) \neq 0$ and $\mu(j) \geq_j i$.
(P3) Let $i \in I$ be a man. Then $w_{i,\mu(i)} \leq p_i \leq 1$.
(P4) Let $i \in I$ be a man and $j \in J$ be a woman such that $j \geq_i 0$ and $i \geq_j 0$. Suppose $p_i - \eta > w_{i,j}$. Then $\mu(j) \neq 0$ and $\mu(j) \geq_j i$. Furthermore, if $\mu(j) =_j i$, then $p_{\mu(j)} \geq p_i$.

In [12], a rather complicated charging argument is used to obtain an approximation ratio of $1 + \frac{1}{e}$ by showing that the optimal fractional value of the LP is at most $1 + \frac{1}{e}$ times the size of any matching μ satisfying (P1)–(P4) with respect to some $\mathbf{p}$.

3 Analysis of the Approximation Ratio

In this section, we analyze the approximation ratio of the algorithm of [12] for the case where the maximum tie length is L. Throughout this section, whenever we mention μ and $\mathbf{p}$, we are referring to their values produced by their algorithm. We use $\mathbf{x}$ to refer to the optimal fractional solution of the LP in their algorithm, and we use $\{w_{i,j}\}_{(i,j) \in I \times (J \cup \{0\})}$ and $\{z_{i,j}\}_{(i,j) \in (I \cup \{0\}) \cup J}$ to refer to the auxiliary variables associated with $\mathbf{x}$ as defined in Section 2.2.

3.1 The Charging Argument

Our charging argument is based on an exchange function $h : [0, 1] \times [0, 1] \to \mathbb{R}$ that satisfies the following properties.

(H1) For every $\xi_1, \xi_2 \in [0, 1]$, we have $0 = h(0, \xi_2) \leq h(\xi_1, \xi_2) \leq 1$.
(H2) For every $\xi_1, \xi_2 \in [0, 1]$ such that $\xi_1 > \xi_2$, we have $h(\xi_1, \xi_2) = 1$.
(H3) The function $h(\xi_1, \xi_2)$ is non-decreasing in ξ_1 and non-increasing in ξ_2.

(H4) For every $\xi_1, \xi_2 \in [0,1]$, we have

$$L \cdot \int_{\xi_2 \cdot (1-1/L)}^{\xi_2} \Big(1 - h(\xi_1, \xi)\Big)\, d\xi \leq \max(\xi_2 - \xi_1, 0).$$

Given an exchange function h which satisfies (H1)–(H4), our charging argument is as follows. For every $(i,j) \in I \times J$, we assign to man i a charge of

$$\theta_{i,j} = \int_0^{x_{i,j}} h(1 - p_i, 1 - w_{i,j} - \xi)\, d\xi$$

and to woman j a charge of

$$\phi_{i,j} = \begin{cases} 0 & \text{if } \mu(j) = 0 \text{ or } i >_j \mu(j) \\ x_{i,j} & \text{if } \mu(j) \neq 0 \text{ and } \mu(j) >_j i \\ x_{i,j} - \displaystyle\int_0^{x_{i,j}} h(1 - p_{\mu(j)}, 1 - z_{\mu(j),j} - \xi)\, d\xi & \text{if } \mu(j) \neq 0 \text{ and } \mu(j) =_j i \end{cases}$$

The following lemma shows that the charges are non-negative and cover the value of LP solution.

Lemma 2. *Let $i \in I$ and $j \in J$. Then $\theta_{i,j}$ and $\phi_{i,j}$ satisfy the following conditions.*

(1) $\theta_{i,j} \geq 0$ and $\phi_{i,j} \geq 0$.
(2) $x_{i,j} \leq \theta_{i,j} + \phi_{i,j}$.

Proof. Part (1) is relatively straightforward to establish; see [13] for a proof. We prove part (2) by considering two cases.

Case 1: $p_i \leq w_{i,j}$. Then (H3) implies

$$\begin{aligned} 0 &\leq \int_0^{x_{i,j}} \Big(h(1 - p_i, 1 - w_{i,j} - \xi) - h(1 - p_i, 1 - p_i - \xi)\Big)\, d\xi \\ &= \int_0^{x_{i,j}} \Big(h(1 - p_i, 1 - w_{i,j} - \xi) - 1\Big)\, d\xi \\ &= \theta_{i,j} - x_{i,j} \\ &\leq \theta_{i,j} + \phi_{i,j} - x_{i,j}, \end{aligned}$$

where the first equality follows from (H2), the second equality follows from the definition of $\theta_{i,j}$, and the last inequality follows from part (1).

Case 2: $p_i > w_{i,j}$. We may assume that $x_{i,j} \neq 0$, for otherwise part (1) implies $\theta_{i,j} + \phi_{i,j} \geq 0 = x_{i,j}$. Since $x_{i,j} \neq 0$, constraint (C4) implies $j \geq_i 0$ and $i \geq_j 0$. So (P4) implies $\mu(j) \neq 0$ and $\mu(j) \geq_j i$. We consider two subcases.

Case 2.1: $\mu(j) >_j i$. Then the definition of $\phi_{i,j}$ implies

$$0 = \phi_{i,j} - x_{i,j} \leq \theta_{i,j} + \phi_{i,j} - x_{i,j}$$

where the inequality follows from part (1).

Case 2.2: $\mu(j) =_j i$. Then (P4) implies $p_i \le p_{\mu(j)}$. Also, since $\mu(j) =_j i$, parts (3) and (4) of Lemma 1 imply $z_{\mu(j),j} = z_{i,j} \le w_{i,j}$. Since $p_i \le p_{\mu(j)}$ and $w_{i,j} \ge z_{\mu(j),j}$, (H3) implies

$$\begin{aligned} 0 &\le \int_0^{x_{i,j}} \Big(h(1-p_i, 1-w_{i,j}-\xi) - h(1-p_{\mu(j)}, 1-z_{\mu(j),j}-\xi) \Big)\, d\xi \\ &= \theta_{i,j} + \phi_{i,j} - x_{i,j}, \end{aligned}$$

where the equality follows from the definitions of $\theta_{i,j}$ and $\phi_{i,j}$. □

3.2 Bounding the Charges

To bound the approximation ratio, Lemma 2 implies that it is sufficient to bound the charges. In Lemma 3, we derive an upper bound for the charges incurred by a man using the strict ordering in his preferences. In Lemma 4, we derive an upper bound for the charges incurred by a woman due to indifferences using the bounded tie length assumption. In Lemma 5, we derive an upper bound for the total charges incurred by a matched couple by combining the results of Lemmas 3 and 4.

Lemma 3. *Let $i \in I$ be a man. Then*

$$\sum_{j \in J} \theta_{i,j} \le \int_0^1 h(1-p_i, \xi)\, d\xi.$$

Proof. Let $j_1, \ldots, j_{|J|} \in J$ such that $j_1 >_i j_2 >_i \cdots >_i j_{|J|}$. Then parts (1) and (2) of Lemma 1 imply

$$w_{i,j_k} + x_{i,j_k} \le \begin{cases} w_{i,j_{k+1}} & \text{if } 1 \le k < |J| \\ 1 & \text{if } k = |J| \end{cases} \tag{1}$$

Hence the definitions of $\{\theta_{i,j_k}\}_{1 \le k \le |J|}$ imply

$$\begin{aligned} \theta_{i,j_k} &= \int_0^{x_{i,j_k}} h(1-p_i, 1-w_{i,j_k}-\xi)\, d\xi \\ &= \int_{w_{i,j_k}}^{w_{i,j_k}+x_{i,j_k}} h(1-p_i, 1-\xi)\, d\xi \\ &\le \begin{cases} \displaystyle\int_{w_{i,j_k}}^{w_{i,j_{k+1}}} h(1-p_i, 1-\xi)\, d\xi & \text{if } 1 \le k < |J| \\ \displaystyle\int_{w_{i,j_{|J|}}}^{1} h(1-p_i, 1-\xi)\, d\xi & \text{if } k = |J| \end{cases} \end{aligned}$$

where the inequality follows from (1) and (H1). Thus

$$\begin{aligned}\sum_{j\in J}\theta_{i,j} &= \sum_{1\le k\le |J|}\theta_{i,j_k}\\ &\le \int_{w_{i,j_{|J|}}}^{1} h(1-p_i,1-\xi)\,d\xi + \sum_{1\le k<|J|}\int_{w_{i,j_k}}^{w_{i,j_{k+1}}} h(1-p_i,1-\xi)\,d\xi\\ &= \int_{w_{i,j_1}}^{1} h(1-p_i,1-\xi)\,d\xi\\ &\le \int_0^1 h(1-p_i,1-\xi)\,d\xi\\ &= \int_0^1 h(1-p_i,\xi)\,d\xi,\end{aligned}$$

where the second inequality follows from $w_{i,j_1}\ge 0$ and (H1). □

Lemma 4. *Let $j\in J$ be a woman such that $\mu(j)\ne 0$. Then*

$$\sum_{\substack{i\in I\\ \mu(j)=_j i}} \phi_{i,j} \le \max(p_{\mu(j)} - z_{\mu(j),j}, 0).$$

Proof. Let

$$H(\xi') = \int_{1-z_{\mu(j),j}-\xi'}^{1-z_{\mu(j),j}} \Big(1 - h(1-p_{\mu(j)},\xi)\Big)\,d\xi$$

for every $\xi'\in[0,1]$. Then (H1) and (H3) imply that H is concave and non-decreasing. Also (H4) implies

$$\begin{aligned} L\cdot H\Big(\frac{1-z_{\mu(j),j}}{L}\Big) &= L\cdot\int_{(1-z_{\mu(j),j})(1-1/L)}^{1-z_{\mu(j),j}} \Big(1-h(1-p_{\mu(j)},\xi)\Big)\,d\xi\\ &\le \max(p_{\mu(j)} - z_{\mu(j),j}, 0). \end{aligned} \tag{2}$$

Let $I' = \{i\in I\colon \mu(j) =_j i\}$. Then $|I'|\le L$ since L is the maximum tie-length. Let $i_1,\dots,i_{|I'|}\in I$ such that $I' = \{i_1,\dots,i_{|I'|}\}$. Let

$$\xi_k = \begin{cases} x_{i_k,j} & \text{if } 1\le k\le |I'|\\ 0 & \text{if } |I'| < k \le L\end{cases}$$

Then the definition of $z_{\mu(j),j}$ implies

$$1 - z_{\mu(j),j} = 1 - \sum_{\substack{i\in I\\ \mu(j)>_j i}} x_{i,j} \ge \sum_{i\in I} x_{i,j} - \sum_{\substack{i\in I\\ \mu(j)>_j i}} x_{i,j} \ge \sum_{\substack{i\in I\\ \mu(j)=_j i}} x_{i,j} = \sum_{1\le k\le L}\xi_k,$$

where the first inequality follows from constraint (C2), and the second equality follows from the definitions of $\{\xi_k\}_{1\le k\le L}$. Hence the monotonicity and concavity of H imply

$$L \cdot H\Big(\frac{1-z_{\mu(j),j}}{L}\Big) \ge L \cdot H\Big(\frac{1}{L}\sum_{1\le k\le L}\xi_k\Big) \ge \sum_{1\le k\le L} H(\xi_k). \tag{3}$$

Thus the definitions of $\{\phi_{i,j}\}_{i\in I}$ imply

$$\begin{aligned}
\sum_{\substack{i\in I\\ \mu(j)=_j i}} \phi_{i,j} &= \sum_{\substack{i\in I\\ \mu(j)=_j i}} \Big(x_{i,j} - \int_0^{x_{i,j}} h(1-p_{\mu(j)}, 1-z_{\mu(j),j}-\xi)\,d\xi\Big)\\
&= \sum_{\substack{i\in I\\ \mu(j)=_j i}} \int_{1-z_{\mu(j),j}-x_{i,j}}^{1-z_{\mu(j),j}} \Big(1-h(1-p_{\mu(j)},\xi)\Big)\,d\xi\\
&= \sum_{\substack{i\in I\\ \mu(j)=_j i}} H(x_{i,j})\\
&= \sum_{1\le k\le L} H(\xi_k)\\
&\le L\cdot H\Big(\frac{1-z_{\mu(j),j}}{L}\Big)\\
&\le \max(p_{\mu(j)} - z_{\mu(j),j}, 0),
\end{aligned}$$

where the third equality follows from the definition of H, the fourth equality follows from the definitions of $\{\xi_k\}_{1\le k\le L}$, the first inequality follows from (3), and the second inequality follows from (2). □

Lemma 5. *Let $i \in I$ and $j \in J \cup \{0\}$ such that $\mu(i) = j$. Then the following conditions hold.*

(1) If $j \neq 0$, then

$$\sum_{j'\in J}\theta_{i,j'} + \sum_{i'\in I}\phi_{i',j} \le 1 + \int_{1-p_i}^{1} h(1-p_i,\xi)\,d\xi.$$

(2) If $j = 0$, then $\theta_{i,j'} = 0$ for every $j' \in J$.

Proof.

(1) Suppose $j \neq 0$. Then (P1) implies $j \ge_i 0$ and $i \ge_j 0$. So part (4) of Lemma 1 implies

$$z_{i,j} \le w_{i,j} \le p_i,$$

where the second inequality follows from (P3). So the definitions of $\{\phi_{i',j}\}_{i'\in I}$ imply

$$\sum_{i'\in I}\phi_{i',j} = \sum_{\substack{i'\in I\\ \mu(j)=_j i'}} \phi_{i',j} + \sum_{\substack{i'\in I\\ \mu(j)>_j i'}} x_{i',j} \le \max(p_i - z_{i,j}, 0) + z_{i,j} = p_i, \tag{4}$$

where the first inequality follows from Lemma 4 and the definition of $z_{i,j}$, and the last equality follows from $p_i \geq z_{i,j}$. Also, by Lemma 3, we have

$$\begin{aligned}
\sum_{j' \in J} \theta_{i,j'} &\leq \int_0^1 h(1-p_i, \xi)\, d\xi \\
&= \int_0^{1-p_i} h(1-p_i, \xi)\, d\xi + \int_{1-p_i}^1 h(1-p_i, \xi)\, d\xi \\
&= \int_0^{1-p_i} 1\, d\xi + \int_{1-p_i}^1 h(1-p_i, \xi)\, d\xi \\
&= 1 - p_i + \int_{1-p_i}^1 h(1-p_i, \xi)\, d\xi, \qquad (5)
\end{aligned}$$

where the second equality follows from (H2). Combining (4) and (5) gives the desired inequality.

(2) Suppose $j = 0$. Let $j' \in J$. Since $\mu(i) = j = 0$, (P3) implies

$$1 \geq p_i \geq w_{i,0} = 1,$$

where the last equality follows from the definition of $w_{i,0}$. Hence the definition of $\theta_{i,j'}$ implies

$$\theta_{i,j'} = \int_0^{x_{i,j'}} h(1-p_i, 1-w_{i,j'}-\xi)\, d\xi = \int_0^{x_{i,j'}} h(0, 1-w_{i,j'}-\xi)\, d\xi = 0,$$

where the second equality follows from $p_i = 1$, and the third equality follows from (H1). □

3.3 The Approximation Ratio

To obtain the approximation ratio, it remains to pick a good exchange function h satisfying (H1)–(H4) such that the right hand side of part (1) of Lemma 5 is small. Using a similar technique as in [12], we can formulate this as an infinite-dimensional factor-revealing linear program. More specifically, we can minimize

$$\sup_{\xi_1 \in [0,1]} \int_{\xi_1}^1 h(\xi_1, \xi)\, d\xi$$

over the set of all functions h which satisfies (H1)–(H4). Notice that the objective value and the constraints induced by (H1)–(H4) are linear in h. However, the space of all feasible solutions is infinite-dimensional. One possible approach to the infinite-dimensional factor-revealing linear program is to obtaining a numerical solution via a suitable discretization. Using the numerical results as guidance, we obtain the candidate exchange function

$$h(\xi_1, \xi_2) = \max\left(\{0\} \cup \left\{(1-\tfrac{1}{L})^k : k \in \{0,1,2,\dots\} \text{ and } \xi_1 > \xi_2 \cdot (1-\tfrac{1}{L})^k\right\}\right). \qquad (6)$$

The following lemma provides a formal analytical proof that it satisfies (H1)–(H4) and achieves an objective value of $(1-\frac{1}{L})^L$.

Lemma 6. *Let h be the function defined by (6). Then the following conditions hold.*

(1) The function h satisfies (H1)–(H4).

(2) For every $\xi_1 \in [0,1]$, we have $\displaystyle\int_{\xi_1}^{1} h(\xi_1,\xi)\,d\xi \le \Big(1-\frac{1}{L}\Big)^L$.

Proof.

(1) It is straightforward to see that (H1)–(H3) hold by inspecting the definition of h. To show that (H4) holds, let $\xi_1,\xi_2 \in [0,1]$. We consider three cases.
Case 1: $\xi_2 \le \xi_1$. Then

$$L\cdot\int_{\xi_2\cdot(1-1/L)}^{\xi_2}\Big(1-h(\xi_1,\xi)\Big)\,d\xi = L\cdot\int_{\xi_2\cdot(1-1/L)}^{\xi_2}(1-1)\,d\xi = 0 \\ = \max(\xi_2-\xi_1,0).$$

Case 2: $\xi_2 > \xi_1 = 0$. Then

$$L\cdot\int_{\xi_2\cdot(1-1/L)}^{\xi_2}\Big(1-h(\xi_1,\xi)\Big)\,d\xi = L\cdot\int_{\xi_2\cdot(1-1/L)}^{\xi_2}(1-0)\,d\xi = \xi_2 \\ = \max(\xi_2-\xi_1,0).$$

Case 3: $\xi_2 > \xi_1 > 0$. Let $k \in \{0,1,2,\dots\}$ such that $(1-\frac{1}{L})^{k+1} < \frac{\xi_1}{\xi_2} \le (1-\frac{1}{L})^k$. Then

$$\begin{aligned}
&L\cdot\int_{(1-1/L)\cdot\xi_2}^{\xi_2}\Big(1-h(\xi_1,\xi)\,d\xi\Big)\\
&=\xi_2 - L\cdot\int_{(1-1/L)\cdot\xi_2}^{\xi_2} h(\xi_1,\xi)\,d\xi\\
&=\xi_2 - L\cdot\int_{(1-1/L)\cdot\xi_2}^{\xi_1/(1-1/L)^k} h(\xi_1,\xi)\,d\xi - L\cdot\int_{\xi_1/(1-1/L)^k}^{\xi_2} h(\xi_1,\xi)\,d\xi\\
&=\xi_2 - L\cdot\int_{(1-1/L)\cdot\xi_2}^{\xi_1/(1-1/L)^k}\Big(1-\frac{1}{L}\Big)^k d\xi - L\cdot\int_{\xi_1/(1-1/L)^k}^{\xi_2}\Big(1-\frac{1}{L}\Big)^{k+1} d\xi\\
&=\xi_2 - L\cdot(\xi_1-\xi_2\cdot(1-\tfrac{1}{L})^{k+1}) - L\cdot(\xi_2\cdot(1-\tfrac{1}{L})^{k+1}-\xi_1\cdot(1-\tfrac{1}{L}))\\
&=\xi_2-\xi_1\\
&=\max(\xi_2-\xi_1,0).
\end{aligned}$$

(2) Let $\xi_1 \in [0,1]$. We may assume that $\xi_1 > 0$, for otherwise

$$\int_{\xi_1}^{1} h(\xi_1,\xi)\,d\xi = \int_{\xi_1}^{1} 0\,d\xi = 0 \le \Big(1-\frac{1}{L}\Big)^L.$$

Let $k \in \{0, 1, 2, \ldots\}$ such that $(1 - \frac{1}{L})^{k+1} < \xi_1 \leq (1 - \frac{1}{L})^k$. Then

$$\begin{aligned}
&\int_{\xi_1}^{1} h(\xi_1, \xi)\, d\xi \\
&= \int_{\xi_1/(1-1/L)^k}^{1} h(\xi_1, \xi)\, d\xi + \sum_{0 \leq k' < k} \int_{\xi_1/(1-1/L)^{k'}}^{\xi_1/(1-1/L)^{k'+1}} h(\xi_1, \xi)\, d\xi \\
&= \int_{\xi_1/(1-1/L)^k}^{1} \left(1 - \frac{1}{L}\right)^{k+1} d\xi + \sum_{0 \leq k' < k} \int_{\xi_1/(1-1/L)^{k'}}^{\xi_1/(1-1/L)^{k'+1}} \left(1 - \frac{1}{L}\right)^{k'+1} d\xi \\
&= \left(\left(1 - \frac{1}{L}\right)^{k+1} - \xi_1 \cdot \left(1 - \frac{1}{L}\right)\right) + \sum_{0 \leq k' < k} \frac{\xi_1}{L} \\
&= (1 - \tfrac{1}{L})^{k+1} + \tfrac{\xi_1}{L}(k - L + 1). \qquad (7)
\end{aligned}$$

We consider three cases.
Case 1: $k = L - 1$. Then (7) implies

$$\int_{\xi_1}^{1} h(\xi_1, \xi)\, d\xi = (1 - \tfrac{1}{L})^{k+1} + \tfrac{\xi_1}{L}(k - L + 1) = (1 - \tfrac{1}{L})^L.$$

Case 2: $k \geq L$. Then (7) implies

$$\begin{aligned}
\int_{\xi_1}^{1} h(\xi_1, \xi)\, d\xi &= (1 - \tfrac{1}{L})^{k+1} + \tfrac{\xi_1}{L}(k - L + 1) \\
&\leq (1 - \tfrac{1}{L})^{k+1} + \tfrac{1}{L}(k - L + 1)(1 - \tfrac{1}{L})^k \\
&= (1 - \tfrac{1}{L})^L \cdot \tfrac{k}{L} \cdot (1 - \tfrac{1}{L})^{k-L} \\
&\leq (1 - \tfrac{1}{L})^L \cdot e^{k/L-1} \cdot e^{(L-k)/L} \\
&= (1 - \tfrac{1}{L})^L,
\end{aligned}$$

where the first inequality follows from $\xi_1 \leq (1 - \frac{1}{L})^k$, and the second inequality follows from $e^{k/L-1} \geq \frac{k}{L}$ and $e^{-1/L} \geq 1 - \frac{1}{L}$.
Case 3: $k \leq L - 2$. Then (7) implies

$$\begin{aligned}
\int_{\xi_1}^{1} h(\xi_1, \xi)\, d\xi &= (1 - \tfrac{1}{L})^{k+1} + \tfrac{\xi_1}{L}(k - L + 1) \\
&< (1 - \tfrac{1}{L})^{k+1} - \tfrac{1}{L}(L - k - 1)(1 - \tfrac{1}{L})^{k+1} \\
&= (1 - \tfrac{1}{L})^L \cdot \tfrac{k+1}{L-1} \cdot (1 + \tfrac{1}{L-1})^{L-k-2} \\
&\leq (1 - \tfrac{1}{L})^L \cdot e^{(k+1)/(L-1)-1} \cdot e^{(L-k-2)/(L-1)} \\
&= (1 - \tfrac{1}{L})^L,
\end{aligned}$$

where the first inequality follows from $\xi_1 > (1 - \frac{1}{L})^{k+1}$, and the second inequality follows from $e^{(k+1)/(L-1)-1} \geq \frac{k+1}{L-1}$ and $e^{1/(L-1)} \geq 1 + \frac{1}{L-1}$. □

Lemma 7. $\sum_{(i,j)\in I\times J} x_{i,j} \le \Big(1+\Big(1-\frac{1}{L}\Big)^L\Big)\cdot|\mu|.$

Proof. Consider the charging argument with the exchange function h as defined by (6). By part (1) of Lemma 6, the function h satisfies (H1)–(H4). Lemma 2 implies

$$\begin{aligned}\sum_{(i,j)\in I\times J} x_{i,j} &\le \sum_{(i,j)\in I\times J}(\theta_{i,j}+\phi_{i,j})\\ &= \sum_{(i,j)\in\mu}\Big(\sum_{j'\in J}\theta_{i,j'}+\sum_{i'\in I}\phi_{i',j}\Big)+\sum_{\substack{i\in I\\ \mu(i)=0}}\sum_{j\in J}\theta_{i,j}+\sum_{\substack{j\in J\\ \mu(j)=0}}\sum_{i\in I}\phi_{i,j}.\end{aligned} \tag{8}$$

Part (1) of Lemma 5 implies

$$\begin{aligned}\sum_{(i,j)\in\mu}\Big(\sum_{j'\in J}\theta_{i,j'}+\sum_{i'\in I}\phi_{i',j}\Big) &\le \sum_{(i,j)\in\mu}\Big(1+\int_{1-p_i}^{1} h(1-p_i,\xi)\,d\xi\Big)\\ &\le \sum_{(i,j)\in\mu}\Big(1+\Big(1-\frac{1}{L}\Big)^L\Big)\\ &= (1+(1-\tfrac{1}{L})^L)\cdot|\mu|,\end{aligned} \tag{9}$$

where the second inequality follows from part (2) of Lemma 6. Part (2) of Lemma 5 implies

$$\sum_{\substack{i\in I\\ \mu(i)=0}}\sum_{j\in J}\theta_{i,j}=0. \tag{10}$$

The definitions of $\{\phi_{i,j}\}_{(i,j)\in I\times J}$ imply

$$\sum_{\substack{j\in J\\ \mu(j)=0}}\sum_{i\in I}\phi_{i,j}=0. \tag{11}$$

Combining (8)–(11) gives the desired inequality. □

Using Lemma 7, it is straightforward to establish the following two theorems; see [13] for proof details.

Theorem 1. *There exists a $(1+(1-\frac{1}{L})^L)$-approximation algorithm for the maximum stable matching problem with one-sided ties and incomplete lists where the maximum tie length is L.*

Theorem 2. *For the maximum stable matching problem with one-sided ties where the maximum tie length is L, the integrality gap of the LP formulation in Section 2.2 is $1+(1-\frac{1}{L})^L$.*

References

1. Bauckholt, F., Pashkovich, K., Sanità, L.: On the approximability of the stable marriage problem with one-sided ties (2018), arXiv:1805.05391
2. Chiang, R., Pashkovich, K.: On the approximability of the stable matching problem with ties of size two (2019), arXiv:1808.04510
3. Dean, B.C., Jalasutram, R.: Factor revealing LPs and stable matching with ties and incomplete lists. In: Proceedings of the 3rd International Workshop on Matching Under Preferences. pp. 42–53 (2015)
4. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. American Mathematical Monthly **69**(1), 9–15 (1962)
5. Gale, D., Sotomayor, M.A.O.: Some remarks on the stable matching problem. Discrete Applied Mathematics **11**(3), 223–232 (1985)
6. Halldórsson, M.M., Iwama, K., Miyazaki, S., Yanagisawa, H.: Randomized approximation of the stable marriage problem. Theoretical Computer Science **325**(3), 439–465 (2004)
7. Halldórsson, M.M., Iwama, K., Miyazaki, S., Yanagisawa, H.: Improved approximation results for the stable marriage problem. ACM Transactions on Algorithms **3**(3), 30 (2007)
8. Huang, C.C., Kavitha, T.: Improved approximation algorithms for two variants of the stable marriage problem with ties. Mathematical Programming **154**(1), 353–380 (2015)
9. Irving, R.W.: Stable marriage and indifference. Discrete Applied Mathematics **48**(3), 261–272 (1994)
10. Iwama, K., Miyazaki, S., Yanagisawa, H.: A 25/17-approximation algorithm for the stable marriage problem with one-sided ties. Algorithmica **68**(3), 758–775 (2014)
11. Király, Z.: Linear time local approximation algorithm for maximum stable marriage. Algorithms **6**(3), 471–484 (2013)
12. Lam, C.K., Plaxton, C.G.: A $(1 + 1/e)$-approximation algorithm for maximum stable matching with one-sided ties and incomplete lists. In: Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 2823–2840 (2019)
13. Lam, C.K., Plaxton, C.G.: Maximum stable matching with one-sided ties of bounded length. Tech. Rep. TR–19–03, Department of Computer Science, University of Texas at Austin (July 2019)
14. McDermid, E.J.: A 3/2-approximation algorithm for general stable marriage. In: Proceedings of the 36th International Colloquium on Automata, Languages and Programming. pp. 689–700 (2009)
15. Paluch, K.: Faster and simpler approximation of stable matchings. Algorithms **7**(2), 189–202 (2014)
16. Radnai, A.: Approximation algorithms for the stable marriage problem. Master's thesis, Department of Computer Science, Eötvös Loránd University (2014)
17. Roth, A.E.: The evolution of the labor market for medical interns and residents: A case study in game theory. Journal of Political Economy **92**(6), 991–1016 (1984)
18. Rothblum, U.G.: Characterization of stable matchings as extreme points of a polytope. Mathematical Programming **54**(1), 57–67 (1992)
19. Vande Vate, J.H.: Linear programming brings marital bliss. Operations Research Letters **8**(3), 147–153 (1989)
20. Yanagisawa, H.: Approximation algorithms for stable marriage problems. Ph.D. thesis, Graduate School of Informatics, Kyoto University (2007)