Beyond Stability: Improved Efficiency Guarantees for α-Stable Matchings

Isabel Fernandez Abad[1] , **Sophie Klumper**[1] and **Guido Schäfer**[1,2]
[1]Centrum Wiskunde & Informatica (CWI), The Netherlands
[2]Institute for Logic, Language and Computation (ILLC), University of Amsterdam, The Netherlands
isabel.fernandez.abad@gmail.com, {s.j.klumper, g.schaefer}@cwi.nl

Abstract

Stable matching mechanisms are fundamental to market design but face an inherent tension between stability and social welfare optimality. We study a natural relaxation of stability, termed α-*stability*, which models agents as willing to deviate only when the potential improvement is sufficiently large. Under α-stability, no pair of agents can deviate and improve their valuations by more than a factor of $1/\alpha$, with $\alpha \in (0, 1]$. We provide a complete characterization of the stability–efficiency tradeoff under asymmetric valuations. This tradeoff depends on the degree of asymmetry $\mu \in (0, 1]$, which bounds the ratio between agents' valuations for any pair. Our results show that relaxing stability can substantially improve achievable efficiency guarantees. We further present a polynomial-time algorithm that computes an α-stable matching attaining the best possible efficiency guarantee. For $\alpha \leq \mu/(\mu + 1)$, our algorithm achieves 1-efficiency; for larger α, it computes an α-stable matching achieving at least $(1/\alpha) \cdot \mu/(\mu+1)$ of the optimal social welfare. Remarkably, our algorithm inflates the values of an optimal matching and then applies the Gale–Shapley algorithm to the modified instance. Finally, we show that computing an optimal α-stable matching is NP-hard, even under slight relaxations of stability, i.e., for α close to 1.

1 Introduction

Matching theory studies how to assign agents to each other based on their preferences in a way that satisfies certain desirable properties. One of the most extensively studied problems in this field is the *stable marriage problem* [Gale and Shapley, 1962; Gusfield and Irving, 1989; Knuth, 1997; Roth and Sotomayor, 1990], which involves two sets of agents L and R of size $n \geq 1$ each, where each agent ranks all members of the opposite set in a strict order of preference. A matching is said to be *stable* if there is no pair of agents who would both prefer to be matched with each other over their current matches. Stable matchings have been a cornerstone of market design with wide-ranging applications in assigning students to schools, medical graduates to residency programs, employees to projects, or students to courses and seminars.

In this work, we study a more general variant of the classical stable matching problem, where agents have *cardinal preferences* over the members of the opposite group, rather than ordinal ones. Such cardinal preferences do not only capture who is preferred over whom, but also quantify by how much. Additionally, we consider cardinal preferences that can encode *indifferences* and *incompatibilities* among the agents. Each agent's cardinal preferences are defined by numerical values (called *valuations*) specifying how much the agent would like to be matched to another agent. More specifically, we use v_{ij} to denote the (positive) valuation that an agent $i \in L$ assigns to $j \in R$, and w_{ji} to denote the (positive) valuation that agent $j \in R$ assigns to $i \in L$. This variant of the stable matching problem is also known as the *stable matching problem with cardinal preferences, ties allowed, and incomplete preference lists* (or, *SMCTI* for short). Stable matchings are guaranteed to exist for SMCTI (see, e.g., [Gusfield and Irving, 1989]).

We are interested in the tradeoff between stability and the social welfare efficiency. Given a matching M, the *social welfare* refers to the total valuations of all matched pairs, i.e., $SW(M) = \sum_{(i,j)\in M}(v_{ij} + w_{ji})$. A stable matching M is said to be γ-*efficient* with $\gamma \in (0, 1]$ if it recovers a γ-fraction of the optimal social welfare (i.e., the maximum social welfare over *all* matchings). Ideally, we would like to ensure that a stable matching provides a strong efficiency guarantee. However, imposing stability inevitably leads to a significant loss in social welfare efficiency (see the example below). As it turns out, this loss depends on the *degree of asymmetry* in the agents' valuations. We capture this by introducing an asymmetry parameter $\mu \in (0, 1]$: the valuations are μ-*bounded* if $v_{ij}/w_{ji} \in [\mu, 1/\mu]$ for every pair (i, j) with $v_{ij} > 0$ and $w_{ji} > 0$. Note that the valuations are symmetric for $\mu = 1$, and become increasingly asymmetric as μ decreases.

Consider the example in Figure 1 and note that the valuations are μ-bounded. It is easy to verify that, for $\epsilon > 0$, the instance has a unique stable matching $M = \{(i_2, j_1)\}$. On the other hand, the social welfare optimal matching is $M^* = \{(i_1, j_1), (i_2, j_2)\}$. In terms of social welfare, we have $SW(M) = 1 + \epsilon$ and $SW(M^*) = (\mu+1)/\mu$. Thus, the efficiency of the (unique) stable matching is $(1+\epsilon)\mu/(\mu+1)$. As a consequence, for this instance we cannot guarantee that

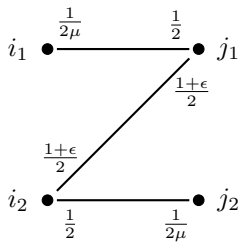

Figure 1: Example with $n = 2$ agents and valuations (v_{ij}) and (w_{ji}) shown as left and right edge labels, respectively.

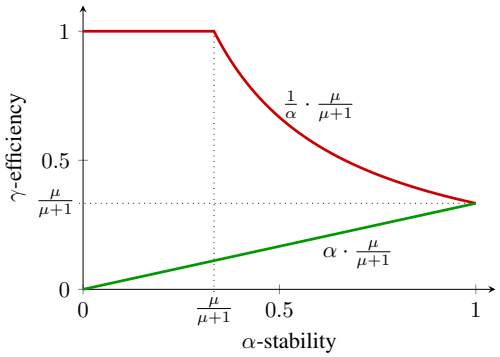

Figure 2: Tradeoff between α-stability and γ-efficiency for $\mu = \frac{1}{2}$. Lower curve (green): lower bound on the efficiency of any α-stable matching. Upper curve (red): upper bound on the best-possible efficiency guarantee of α-stable matchings.

any stable matching recovers (strictly) more than a fraction of $\mu/(\mu+1)$ of the optimal social welfare. In particular, the efficiency of a stable matching becomes unbounded as μ approaches 0, or, said differently, as the valuations become increasingly asymmetric. In fact, this tradeoff has been observed before by Anshelevich *et al.* [2013] for the two limiting cases $\mu = 1$ and $\mu \to 0$; but, here, we get a smooth interpolation for any μ between these extreme points—which we show is tight also (see our contributions below).

A further drawback of the classical notion of stability is that it assumes agents will always deviate, even if the respective gain is vanishingly small. This is an overly strong assumption in settings where agents may be reluctant to deviate. As an example, consider the situation of deciding on money transfers between different accounts, where a commission discourages moves unless the gain justifies the expense. Another example are housing markets, where the decision to move to a better home must also account for the financial and personal costs of relocating. In this context, an approach that has been pursued in the literature is to relax the stability notion by incorporating a fixed, additive *switching cost* [Anshelevich *et al.*, 2013]. While an additive switching cost may be appropriate in some scenarios, in others a cost that scales proportionally with the valuations may be more realistic.

In this work, we build on a relaxed stability notion, termed α-*stability*, introduced by Anshelevich *et al.* [2013]. The key idea is to model agents as willing to deviate only when the improvement is sufficiently large, with the required threshold being determined by a parameter α. Formally, a matching M is α-*stable* with $\alpha \in (0, 1]$ if there is no pair (i, j) such that $v_{iM(i)} < \alpha v_{ij}$ and $w_{jM(j)} < \alpha w_{ji}$, where $M(i)$ and $M(j)$ denote the matches of i and j under M, respectively. As α decreases, agents require a larger improvement to switch, making the stability condition less restrictive.

As mentioned above, α-stability was first introduced in [Anshelevich *et al.*, 2013], where they restrict attention to the case of symmetric valuations (i.e., $\mu = 1$). To the best of our knowledge, we are the first to study the more complex case of asymmetric valuations that are μ-bounded. The idea of studying relaxation of fundamental solution concepts has also been explored in adjacent fields, such as *approximate Nash equilibria* by Daskalakis *et al.* [2009] and relaxed fairness notions such as *envy-freeness up to one good (EF1)* by Budish [2011].

1.1 Our Contributions

We study how the relaxed notion of α-stability influences social welfare efficiency under asymmetric (μ-bounded) valuations, which we term the *stability–efficiency tradeoff*.

Our main results are as follows:

1. *Characterization of Stability–Efficiency Tradoff:* We provide a complete characterization of the stability–efficiency tradeoff (Section 3); see Figure 2 for an illustration. We prove that any α-stable matching is at least $(\alpha \cdot \mu/(\mu+1))$-efficient (Proposition 2). Thus, the worst-case efficiency of α-stable matchings becomes worse as α decreases. On the other hand, we prove an upper bound of $1/\alpha \cdot \mu/(\mu+1)$ on the best-case efficiency of α-stable matchings. In particular, for smaller values of α, α-stable matchings can achieve a (significantly) better efficiency guarantee than stable matchings, with the magnitude of improvement depending on the ratio $\mu/(\mu+1)$. Also, for $\alpha \leq \mu/(\mu+1)$, any optimal matching is α-stable, implying that α-stable matchings can be 1-efficient. At the same time, our result shows that it is impossible to surpass the bound of $\min(1, 1/\alpha \cdot \mu/(\mu+1))$. This provides a complete characterization of the stability–efficiency tradeoff for α-stable matchings under μ-bounded valuations; previously, only the case $\mu = 1$ was understood [Anshelevich *et al.*, 2013].

2. *Computing α-Stable Matchings:* We present a polynomial-time algorithm, termed α-*Boosting*, that computes an α-stable matching achieving the best-possible worst-case efficiency guarantee (Section 4, Theorem 1). For $\alpha \leq \mu/(\mu+1)$, the algorithm simply returns a social welfare optimal matching, which is α-stable and achieves a 1-approximation. For $\alpha > \mu/(\mu+1)$, α-Boosting uses a more sophisticated approach inspired by [Colini-Baldeschi *et al.*, 2024]: it first computes a social welfare optimal matching, inflates the values of its edges by a factor $1/\alpha$, and then applies the Gale–Shapley algorithm to the modified instance. Remarkably, we can show that this algorithm has an approximation guarantee of $1/\alpha \cdot \mu/(\mu+1)$, which

is best possible.

3. *Hardness of Optimal α-Stable Matchings:* We consider the problem of computing an α-stable matching of maximum social welfare (Section 5). We derive two integer linear programming formulations that characterize the set of α-stable matchings. For the stable matching problem with strict preference orders, Vande Vate [1989] and Rothblum [1992] showed that the corresponding linear programming relaxations are integral. However, such a result seems elusive for α-SMCTI. We prove that for any $\alpha \in ({}^{n-1}/_{n}, 1)$, deciding whether an instance of α-SMCTI admits an α-stable matching achieving at least a given social welfare is NP-complete (Theorem 2). In particular, this implies that computing an optimal α-stable matching is NP-hard for this range of α.

1.2 Related Work

The stable matching problem was first introduced by Gale and Shapley [1962], together with an algorithm that computes a stable matching which is optimal in terms of social welfare for the proposing side. Their model assumes that (i) preference lists are complete, and (ii) there are no ties in the preferences. Gale and Sotomayor [1985] considered the case of incomplete preferences and studied how changing the length on an agent's preference list affects their welfare in the final matching.

Later work has explored generalizations of the problem that drop both of the assumptions described above. Iwama *et al.* [1999] showed that when both assumptions are removed, the problem of determining whether a (perfect) stable matching exists is NP-hard. When only (ii) is dropped, Manlove *et al.* [2002] proved that finding an optimal stable matching is NP-hard, and Erdil and Ergin [2017] showed that the matching resulting from the Gale–Shapley algorithm may not be optimal for either side. More recently, Haas [2021] proposed heuristic methods to improve the social welfare of these matchings when ties occur.

A different way to generalize the stable marriage problem is to consider weighted preferences. Anshelevich *et al.* [2013] do this and define the (multiplicative) notion of α-stability, which we consider in this paper. They consider three scenarios: vertex-labeled graphs, symmetric edge-labeled graphs, and asymmetric edge-labeled graphs.[1] In the latter case, they showed that the social welfare of an α-stable matching can be arbitrarily worse than that of an optimal matching, but did not explore this setting further. For symmetric edge-labeled graphs, however, they proved that there always exists an α-stable matching with social welfare at least ${}^{1}/_{2\alpha}$ times that of an optimal matching.[2] They also provided an algorithm that finds such a matching in $O(n^2)$ time.

α-Stability was also analyzed in the context of fractional matchings by Caragiannis *et al.* [2021]. They propose a polynomial time algorithm that computes an α-stable fractional matching which guarantees an α-approximation. They also prove that the problem of finding an optimal α-stable fractional matching is hard to approximate. Different notions of stability have also been considered in the literature (e.g., [Chen *et al.*, 2021; Echenique *et al.*, 2025; Lin *et al.*, 2024; Pini *et al.*, 2013; Troyan *et al.*, 2020]).

The linear programming approach to the stable matching problem has been widely studied. Vande Vate [1989] was the first to show that the polytope defined by the stable matching constraints has integral extreme points, thereby proving that an optimal stable matching can be found in polynomial time. Rothblum [1992] later expanded on these results and offered simplified proofs. However, they do not easily generalize to the α-stability setting. On the other hand, Manlove *et al.* [2002] considered a series of harder variants of the problem, which they showed to be NP-hard.

Finally, the stable marriage problem can be seen as a special case of the Generalized Assignment Problem for unit sizes and capacities. Colini-Baldeschi *et al.* [2024] study different variants of this problem when one group is strategic in a learning-augmented environment. In particular, they introduce a mechanism, called *Boost*, which increases the values of edges in the predicted matching such that these edges are more likely to be matched during a modified run of the Gale–Shapley algorithm. Our α-Boosting algorithm builds upon this procedure.

2 Preliminaries

Stable Matching with Cardinal Preferences. An instance of the *stable matching problem with cardinal preferences* is defined as follows: We are given two sets L and R, each consisting of n agents.[3] We use i and j to refer to agents from the sets L and R, respectively. Each agent has cardinal preferences (allowing for indifferences and incompatibilities) over the agents in the other set. More specifically, each agent $i \in L$ has a non-negative *valuation* v_{ij} for every $j \in R$, and each agent $j \in R$ has a non-negative *valuation* w_{ji} for every $i \in L$. The valuations of an agent comprise their *preference list*. The interpretation of the valuations (v_{ij}) for agent $i \in L$ is as follows: agent i prefers to be matched to j_1 rather than j_2 if $v_{ij_1} > v_{ij_2}$, i is indifferent between j_1 and j_2 if $v_{ij_1} = v_{ij_2}$, and i does not want to be matched to j if $v_{ij} = 0$. The interpretation of the valuations (w_{ji}) for an agent $j \in R$ follows analogously. We define (i, j) as a *pair* if $i \in L$ and $j \in R$. We say that (i, j) is a *compatible* pair if $v_{ij} \neq 0$ and $w_{ji} \neq 0$. The *preference graph* $\mathrm{PG}(L \cup R, (v_{ij}), (w_{ji}))$ (or simply PG, when clear from context) is the bipartite graph that represents these relational preferences: the vertex set is $L \cup R$, the edge set consists of all compatible pairs, i.e., $E(\mathrm{PG}) = \{(i, j) \in L \times R : v_{ij} \neq 0 \text{ and } w_{ji} \neq 0\}$, and each edge $(i, j) \in E(\mathrm{PG})$ is associated with two (positive) weights v_{ij} and w_{ji}. We assume w.l.o.g. that the preference graph PG is non-empty.

α-Stable Matchings. Given a preference graph PG, a *matching* M_{PG} (or simply M, when clear from context) for

[1] Our μ parameter interpolates between symmetric edge-labeled ($\mu = 1$) and asymmetric edge-labeled graphs ($\mu \to 0$).

[2] Note that in Anshelevich *et al.* [2013], the approximation ratios are presented inversely compared to our notation.

[3] It is w.l.o.g. that $|L| = |R|$, i.e., an instance with $|L| > |R|$ can be reduced to an instance with $|L| = |R'|$ by adding agents to R and defining that $\forall j \in R' \setminus R$ it holds that $v_{ij} = w_{ji} = 0, \forall i \in L$.

PG is a set of edges such that each vertex is incident to at most one edge. Unless specified otherwise, a matching refers to a non-empty matching. We will abuse notation and also denote by M the mapping $L \cup R \to L \cup R \cup \emptyset$ such that $M(i) = j$ and $M(j) = i$ whenever $(i, j) \in M$. Otherwise, if there is no $j \in R$ such that $(i, j) \in M$, we set $M(i) = \emptyset$, in which case we will say that i is *lonely* under M and define $v_{iM(i)} = 0$. For $j \in R$, $M(j) = \emptyset$ and $w_{jM(j)}$ are defined similarly. We use Mat(PG) to refer to the set of matchings for PG. A matching M is said to be *(inclusion-wise) maximal* if there is no edge $e \in E(\mathrm{PG}) \setminus M$ such that $M \cup \{e\}$ is also a matching. A matching is said to be *stable* if no pair strictly prefers each other to their current matches if any, or to being lonely.[4] In this paper we consider a relaxed stability notion, called α-*stability* (see also [Anshelevich *et al.*, 2013]).

Definition 1 (α-Stability)**.** *Let M be a matching for a preference graph* PG*, and let $\alpha \in (0, 1]$. A pair $(i, j) \in E(\mathrm{PG})$ is* α-blocking *if $v_{iM(i)} < \alpha v_{ij}$ and $w_{jM(j)} < \alpha w_{ji}$. M is* α-stable *if there is no α-blocking pair.*[5]

Note that the original stability notion is equivalent to 1-stability. By definition, an α-stable matching M is also α'-stable for every $\alpha' \leq \alpha$. In particular, a stable matching is α-stable for every $\alpha \in (0, 1]$. This observation also implies that, for any preference graph PG, an α-stable matching is guaranteed to exist for every $\alpha \in (0, 1]$ as a stable matching always exists (see [Gusfield and Irving, 1989]). Observe that any α-stable matching must be (inclusion-wise) maximal.

Benchmark. We are interested in how the quality of α-stable matchings changes depending on α. To this end, we define the *social welfare* of a matching M as $SW(M) = \sum_{(i,j)\in M}(v_{ij} + w_{ji})$. That is, the social welfare $SW(M)$ accounts for the sum of the valuations of all pairs in the matching M. Given a preference graph PG, an *optimal matching* is a matching of maximum social welfare and will be denoted by M^*, i.e., $M^* \in \mathrm{Mat}(\mathrm{PG})$ and $SW(M^*) \geq SW(M)$ for every matching $M \in \mathrm{Mat}(\mathrm{PG})$. Throughout this work, we use the optimal social welfare as a benchmark when evaluating the quality of a matching.

Definition 2 (γ-Efficiency)**.** *Let M be a matching and M^* be an optimal matching for a preference graph* PG*. M is said to be* γ-efficient *with $\gamma \in (0, 1]$ if $SW(M) = \gamma SW(M^*)$.*

μ-Bounded Valuations. As it turns out, the efficiency of α-stable matchings does not only depend on the stability parameter α, but also on the "degree of asymmetry" of the underlying valuations. The following notion formalizes this. Given a preference graph PG, the valuations of the agents are μ-*bounded* with $\mu \in (0, 1]$ if $\mu w_{ji} \leq v_{ij} \leq w_{ji}/\mu$ for all $(i, j) \in E(\mathrm{PG})$. Intuitively, for $\mu = 1$ the valuations are symmetric (i.e., $v_{ij} = w_{ji}$ for all $(i, j) \in E(\mathrm{PG})$), and they become increasingly asymmetric as μ decreases.

Computing α-Stable Matchings. We refer to our stable matching problem introduced above as α-*SMCTI* for short, which stands for α-*Stable Matching with Cardinal preferences, Ties allowed, and Incomplete preference lists*. An instance of α-SMCTI is given by a preference graph PG and the parameter α. Given the preference graph PG, we let μ refer to the (largest) μ parameter for which the valuations are μ-bounded (as defined above). We are interested in deriving algorithms that compute α-stable matchings with good efficiency guarantees. In this context, a ρ-*approximation algorithm* with $\rho \in (0, 1]$ is an algorithm that, for any given instance of α-SMCTI, computes in polynomial time an α-stable matching M for PG that is at least ρ-efficient, i.e., $SW(M) \geq \rho SW(M^*)$; ρ is also called the *approximation guarantee*. We also consider the problem of computing an *optimal α-stable matching*, i.e., an α-stable matching of maximum social welfare among all α-stable matchings.

3 Stability–Efficiency Tradeoff

Given an instance of α-SMCTI, our ultimate goal is to compute an α-stable matching with a good (large) efficiency. We begin by analyzing the worst-case and best-case efficiency guarantees achievable by α-stable matchings.

Proposition 1. *Consider an instance of α-SMCTI with preference graph* PG *and μ-bounded valuations. Then*

(a) *if $\alpha \in (0, \frac{\mu}{\mu+1}]$, there is always a 1-efficient α-stable matching, and*

(b) *if $\alpha \in (\frac{\mu}{\mu+1}, 1]$, then for every $\delta > 0$ there exists a preference graph* PG *such that there is no $(\frac{1}{\alpha} \cdot \frac{\mu}{\mu+1} + \delta)$-efficient α-stable matching.*

Proof. Let $\alpha \in (0, \frac{\mu}{\mu+1}]$. We will show that any optimal matching is α-stable. Towards a contradiction, let M^* be an optimal matching that is not α-stable for some $\alpha \in (0, \frac{\mu}{\mu+1}]$. Then there exists an α-blocking pair (i, j), i.e., $v_{iM^*(i)} < \alpha v_{ij}$ and $w_{jM^*(j)} < \alpha w_{ji}$. As the valuations are μ-bounded we obtain

$$v_{iM^*(i)} + w_{M^*(i)i} \leq \tfrac{\mu+1}{\mu} v_{iM^*(i)} < \alpha \cdot \tfrac{\mu+1}{\mu} v_{ij}$$
$$v_{M^*(j)j} + w_{jM^*(j)} \leq \tfrac{\mu+1}{\mu} w_{jM^*(j)} < \alpha \cdot \tfrac{\mu+1}{\mu} w_{ji}.$$

And as $\alpha \cdot \frac{\mu+1}{\mu} \leq 1$, adding the above equations leads to

$$v_{iM^*(i)} + w_{M^*(i)i} + v_{M^*(j)j} + w_{jM^*(j)} < v_{ij} + w_{ji}. \quad (1)$$

The matching $\hat{M} := M^* \setminus \{(i, M^*(i)), (M^*(j), j)\} \cup \{(i, j)\}$ therefore satisfies $SW(\hat{M}) > SW(M^*)$, contradicting optimality of M^*. Observe that the argument also holds if either i or j is lonely under M^*, as (1) remains true, and if i and j are both lonely under M^* then this immediately contradicts optimality of M^*.

Now let $\alpha \in (\frac{\mu}{\mu+1}, 1]$. For each value of $\delta > 0$, we will construct a preference graph such that there is no $(\frac{1}{\alpha} \cdot \frac{\mu}{\mu+1} + \delta)$-efficient α-stable matching. Consider the instance shown in Figure 1, where $\epsilon = \frac{1}{\alpha} - 1 + \epsilon'$ for some $0 < \epsilon' < (1 + \frac{1}{\mu})\delta$ such that $\epsilon \in (\frac{1}{\alpha} - 1, \frac{1}{\mu})$.[6] The only α-stable matching

[4] This is sometimes referred to as *weak stability*.

[5] We impose $\alpha > 0$ as every matching is trivially 0-stable.

[6] This interval is well defined as $\frac{1}{\alpha} - 1 < \frac{1}{\mu}$ since $\alpha > \frac{\mu}{\mu+1}$.

is given by $M = \{(i_2, j_1)\}$, while the optimal matching is $M^* = \{(i_1, j_1), (i_2, j_2)\}$, leading to[7]

$$\frac{SW(M)}{SW(M^*)} = \frac{1+\epsilon}{1+\frac{1}{\mu}} = \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1} + \frac{\epsilon'}{1+\frac{1}{\mu}} < \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1} + \delta.$$

□

Proposition 2. *Given an instance of α-SMCTI with preference graph* PG *and μ-bounded valuations, any α-stable matching is at least $(\alpha \cdot \frac{\mu}{\mu+1})$-efficient.*

Proof. Let M be an α-stable matching, and fix an optimal matching M^*. Due to α-stability, for every $(i,j) \in M^*$, it holds that $v_{iM(i)} \geq \alpha v_{ij}$ or $w_{jM(j)} \geq \alpha w_{ji}$. Thus we can write $M^* = E_1 \cup E_2$, where $E_1 := \{(i,j) \in M^* : v_{iM(i)} \geq \alpha v_{ij}\}$ and $E_2 := \{(i,j) \in M^* : w_{jM(j)} \geq \alpha w_{ji}\}$. In particular,

$$SW(M^*) \leq \sum_{(i,j)\in E_1} (v_{ij} + w_{ji}) + \sum_{(i,j)\in E_2} (v_{ij} + w_{ji}). \quad (2)$$

Consider the first summation in (2). It holds that $w_{ji} \leq \frac{1}{\mu} v_{ij} \leq \frac{1}{\alpha\mu} v_{iM(i)}$, as the valuations are μ-bounded and by definition of E_1. Therefore,

$$\sum_{(i,j)\in E_1} (v_{ij} + w_{ji}) \leq \sum_{(i,j)\in M} \frac{1}{\alpha}\left(1 + \frac{1}{\mu}\right) v_{ij}. \quad (3)$$

We can upper bound the second summation in (2) analogously, and therefore upper bound (2) by $SW(M^*) \leq \frac{1}{\alpha}(1 + \frac{1}{\mu})SW(M)$, proving the claim. □

Note that when $\alpha = 1$, the bounds in Propositions 1 and 2 coincide. Also note that by relaxing the stability requirement to α-stability, the lower bound on the efficiency guarantee decreases whereas the upper bound on the efficiency guarantee increases. We refer to Figure 2 for an illustration of these bounds.

4 Computing α-Stable Matchings

Our aim for this section is to design an algorithm which, given any instance of α-SMCTI with μ-bounded valuations, achieves an approximation guarantee of 1 whenever $\alpha \leq \mu/\mu+1$, and of $1/\alpha \cdot \mu/\mu+1$ otherwise. Note that in light of Proposition 2, this is best possible. We start by introducing our algorithm, called *α-Boosting*, which builds upon [Gale and Shapley, 1962] and [Colini-Baldeschi *et al.*, 2024].

α-Boosting Algorithm. Given a preference graph PG and a parameter $\alpha \in (0, 1]$ as input, α-Boosting first computes an optimal matching M^* for PG. If M^* is α-stable, then M^* is returned. Otherwise, α-Boosting defines an instance PG^α with the same underlying bipartite graph as PG and valuations (v'_{ij}, w'_{ji}) with

$$v'_{ij} = \begin{cases} \frac{1}{\alpha} v_{ij}, & \text{if } (i,j) \in M^*, \\ v_{ij}, & \text{otherwise,} \end{cases}$$

and w'_{ji} defined analogously. It then returns the matching M that is obtained by running the version of the Gale-Shapley algorithm that handles incomplete preferences (see [Gusfield and Irving, 1989]) on PG^α, with ties broken arbitrarily.[8]

Denote $m = |E(\mathrm{PG})|$. Since an optimal matching for PG can be computed in time $O(m^{1+o(1)})$ [Chen *et al.*, 2025], PG^α can be constructed in time $O(m)$, and the Gale-Shapley algorithm on PG^α runs in time $O(m \log n)$, the complexity of α-Boosting is $O(m \log n)$. In the special case of $\mu = 1$, an algorithm with runtime of $O(n^2)$ has already been proposed by Anshelevich *et al.* [2013]. However, it is not trivial whether this algorithm can be generalized for $\mu \in (0, 1)$.

Note that M is a matching for PG if and only if it is a matching for PG^α. Throughout this section, we will specify the preference graph (PG or PG^α) in the subscript when relevant, and omit it otherwise. The main result of this section is the following theorem:

Theorem 1. *Let $\alpha \in (0, 1]$. Then, for any instance of α-SMCTI, α-Boosting returns an α-stable matching and achieves an approximation guarantee of $f(\alpha, \mu)$ with*

$$f(\alpha, \mu) = \begin{cases} 1, & \text{if } \alpha \leq \frac{\mu}{\mu+1}, \\ \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1}, & \text{otherwise,} \end{cases} \quad (4)$$

which is best possible.

Proof. Consider an instance PG of α-SMCTI. If $\alpha \leq \frac{\mu}{\mu+1}$, then α-Boosting returns M^* of PG which is α-stable (Proposition 1), thus achieving an approximation guarantee of 1 in this case. Therefore, assume that $\alpha > \frac{\mu}{\mu+1}$ and let M be the matching returned by α-Boosting. We start by showing that M is α-stable. Let (i,j) be a pair which is not in M. We distinguish two cases:

1. If i did not propose to j during the execution of the Gale-Shapley algorithm, then we must have that $v'_{iM(i)} \geq v'_{ij}$. By definition, it holds that $\frac{1}{\alpha} v_{iM(i)} \geq v'_{iM(i)} \geq v'_{ij} \geq v_{ij}$, so in particular, $v_{iM(i)} \geq \alpha v_{ij}$. Therefore, (i,j) cannot be α-blocking.
2. Otherwise, i proposed to j but j rejected i, so it must hold that $w'_{jM(j)} \geq w'_{ji}$. Again by definition, it holds that $\frac{1}{\alpha} w_{jM(j)} \geq w'_{jM(j)} \geq w'_{ji} \geq w_{ji}$, so $w_{jM(i)} \geq \alpha w_{ji}$ and (i,j) cannot be α-blocking.

Next, we establish the approximation guarantee by fixing an optimal matching M^* for PG and constructing a mapping $g : M^* \to M$. The main idea is that f maps each edge in the optimal matching either to itself or to an adjacent edge of comparable value.

Let $(i,j) \in M^*$. If $(i,j) \in M$, we set $g(i,j) = (i,j)$. Otherwise, $v'_{iM(i)} \geq v'_{ij}$ or $w'_{jM(j)} \geq w'_{ji}$. If the first condition holds, we set $g(i,j) = (i, M(i))$. Since $(i, M(i)) \notin$

[7] Note that $SW(M^*) > 0$, as $E(\mathrm{PG}) \neq \emptyset$.

[8] The Gale–Shapley algorithm operates on ordinal preferences, but any set of cardinal preferences can be uniquely converted into this form.

M^*, it follows that $v_{iM(i)} = v'_{iM(i)} \geq v'_{ij} = \frac{1}{\alpha} v_{ij}$, and combining with the definition of μ-bounded valuations leads to

$$\frac{v_{iM(i)} + w_{M(i)i}}{v_{ij} + w_{ji}} \geq \frac{(1+\mu)v_{iM(i)}}{(1+\frac{1}{\mu})v_{ij}} \geq \frac{\frac{1}{\alpha}(1+\mu)v_{ij}}{(1+\frac{1}{\mu})v_{ij}} = \frac{\mu}{\alpha}.$$

If *only* the second condition holds, we set $g(i,j) = (M(j), j)$, which yields an analogous bound. Now, if $(i,j) \in M$ and $(i,j) \notin M^*$, it could be that two edges $(i,j') \in M^*$ and $(i',j) \in M^*$ are mapped to (i,j). In this case we have that $\alpha v_{ij} \geq v_{ij'}$ and $\alpha w_{ji} \geq w_{ji'}$, so

$$\frac{v_{ij} + w_{ji}}{v_{ij'} + w_{j'i} + v_{i'j} + w_{ji'}} \geq \frac{\frac{1}{\alpha} v_{ij'} + \frac{1}{\alpha} w_{ji'}}{v_{ij'} + w_{j'i} + v_{i'j} + w_{ji'}} \geq \frac{\frac{1}{\alpha} v_{ij'} + \frac{1}{\alpha} w_{ji'}}{(1+\frac{1}{\mu})(v_{ij'} + w_{ji'})} = \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1}.$$

Since $1 > \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1}$, $\frac{\mu}{\alpha} > \frac{1}{\alpha} \cdot \frac{\mu}{\mu+1}$ and edges in M^* are mapped to adjacent edges in M or itself (if in M), we obtain

$$\begin{aligned} \tfrac{1}{\alpha} \cdot \tfrac{\mu}{\mu+1} SW(M^*) &= \tfrac{1}{\alpha} \cdot \tfrac{\mu}{\mu+1} \sum_{(i,j) \in M^*} (v_{ij} + w_{ji}) \\ &\leq \sum_{(i,j) \in M} (v_{ij} + w_{ji}) = SW(M), \end{aligned}$$

proving that α-Boosting is $f(\alpha,\mu)$-approximate as in (4). Furthermore, this is best possible by Proposition 1. □

5 Hardness of Optimal α-Stable Matchings

As shown in the previous section, α-Boosting always returns an α-stable matching that recovers at least $1/\alpha \cdot \mu/(\mu+1)$ of the optimal social welfare. However, it need not be an optimal α-*stable* matching. In this section, we consider the problem of finding an optimal α-stable matching.

Integer Linear Program of α-SMCTI. Given a matching M for a preference graph $\mathrm{PG}(L \cup R, (v_{ij}), (w_{ji}))$, we define its *assignment matrix* as $x(M) = (x_{ij})_{i \in L, j \in R}$, where $x_{ij} = 1$ if the pair $(i,j) \in M$, and $x_{ij} = 0$ otherwise. The set of all matchings can be described by the following system of inequalities:

$$\sum_{j \in R} x_{ij} \leq 1 \qquad \forall i \in L \tag{5}$$

$$\sum_{i \in L} x_{ij} \leq 1 \qquad \forall j \in R \tag{6}$$

$$x_{ij} \in \{0,1\} \quad \forall (i,j) \in E(\mathrm{PG}). \tag{7}$$

Below, we show that the α-stability property can be described by two different linear constraints.

Proposition 3. *Let M be a matching for $\mathrm{PG}(L \cup R, (v_{ij}), (w_{ji}))$ and let $\alpha \in (0,1]$. Then, M is α-stable if and only if $x(M)$ satisfies*

$$\sum_{l : v_{il} \geq \alpha v_{ij}} x_{il} + \sum_{k : w_{jk} \geq \alpha w_{ji}} x_{kj} \geq 1 \quad \forall (i,j) \in E(\mathrm{PG}). \tag{8}$$

Proof. Towards a contradiction, suppose that M is an α-stable matching and there exists $(i,j) \in E(\mathrm{PG})$ such that (8) does not hold. In this case it must be that both summations in (8) are equal to 0, and so $v_{iM(i)} < \alpha v_{ij}$ and $w_{jM(j)} < \alpha w_{ji}$. This implies that (i,j) is an α-blocking pair which contradicts that M is α-stable. For the other direction note that if (8) is satisfied for a matching M, this implies that there is no α-blocking pair $(i,j) \in E(\mathrm{PG})$ as (i,j) would otherwise not satisfy (8), implying that M is α-stable. □

Proposition 4. *Let M be a matching for $\mathrm{PG}(L \cup R, (v_{ij}), (w_{ji}))$ and let $\alpha \in (0,1]$. Then, M is α-stable if and only if $x(M)$ satisfies*

$$\sum_{k : \alpha w_{ji} > w_{jk}} x_{kj} - \sum_{l : v_{il} \geq \alpha v_{ij}} x_{il} \leq 0 \quad \forall (i,j) \in E(\mathrm{PG}). \tag{9}$$

Proof. Let M be an α-stable matching, and suppose that for a certain $(i,j) \in E(\mathrm{PG})$ we have that $\sum_{k : \alpha w_{ji} > w_{jk}} x_{kj} = 1$. This means that j is paired with an agent they find less desirable than i by a factor of α. Since M is α-stable, we must have $v_{iM(i)} \geq \alpha v_{ij}$. In particular, $\sum_{l : v_{il} \geq \alpha v_{ij}} x_{il} = 1$. This settles one direction. For the other direction, we proceed by contrapositive. Suppose that M is not α-stable, and let (i,j) be an α-blocking pair, so $v_{iM(i)} < \alpha v_{ij}$ and $w_{jM(j)} < \alpha w_{ji}$. The first implies that $M(i) \notin \{l : v_{il} \geq \alpha v_{ij}\}$, and therefore $\sum_{l : v_{il} \geq \alpha v_{ij}} x_{il} = 0$. The second implies that $M(j) \in \{k : \alpha w_{ji} > w_{jk}\}$, so $\sum_{k : \alpha w_{ji} > w_{jk}} x_{kj} = 1$ and (9) does not hold for (i,j) which concludes the proof. □

Using Propositions 3 and 4, we can describe the set of all α-stable matchings via constraints (5)–(7), together with either (8) or (9). Thus, computing an α-stable matching is equivalent to solving the integer program that maximizes the objective $\sum_{(i,j) \in E(\mathrm{PG})} (v_{ij} + w_{ji}) x_{ij}$ subject to these constraints. The crux, however, is that integer programs are not known to be solvable in polynomial time.

Notably, for the special case of stable matchings (i.e., $\alpha = 1$) with strict preference orders, Rothblum [1992] and Vande Vate [1989] showed that the extreme points of the polytope defined by (5), (6) and

$$x_{ij} \geq 0 \quad \forall (i,j) \in E(\mathrm{PG}), \tag{10}$$

together with (8) or (9), are integral.[9] Consequently, in this case, solving the corresponding linear program (which can be done in polynomial time) yields an integral optimal solution. That is, for $\alpha = 1$ we can find an optimal α-stable matching by solving the respective linear program. This naturally raises the question of whether the same holds for $\alpha < 1$. We answer this question in the negative by showing that the problem is NP-hard, even when α is arbitrarily close to 1.

To prove our hardness results, we consider certain special cases of our α-SMCTI problem. We denote these more restrictive variants by omitting the respective letters from the acronym. In particular, SMCT (SMT) refers to the stable matching variant with cardinal (ordinal) preferences, ties allowed, and complete preference lists. Given a weak order relation $\succeq_i$ for $i \in L$, we write $j_1 \succeq_i j_2$ to indicate that

[9] In these papers, preference lists are given as order relations, but the generalization to cardinal preferences is almost immediate.

i (weakly) prefers j_1 over j_2; the relation $\succeq_j$ for $j \in R$ is defined analogously.

Given an instance I of SMT and $(i,j) \in L \times R$, we define the *cost* c_{ij} of agent i for j as the position j occupies in i's preference list (the cost c_{ji} of agent j for i is defined analogously). Agents in the same tie group share the same cost, which will be equal to the number of agents ranked strictly above them plus one. We define the *weight* of M as $w(M) := \sum_{u \in L \cup R} c_{uM(u)}$, where $c_{uM(u)} = \infty$ whenever u is lonely under M. We say that a matching is *egalitarian* if its weight is minimum among all stable matchings of I.

Let EGALITARIAN-SMT-OPT denote the optimization problem of finding an egalitarian stable matching given an instance of SMT. Manlove *et al.* [2002] proved that EGALITARIAN-SMT-OPT is hard to approximate. In particular, this means that the following decision problem is also NP-hard:

MIN-COST-SMT: Given an instance of SMT and $K \in \mathbb{R}^+$, is there a stable matching M such that $w(M) \leq K$?

The decision problem we are interested in here is:

MAX-WEIGHT-α-SMCT: Given an instance of α-SMCT and $\omega \in \mathbb{R}^+$, is there an α-stable matching M such that $SW(M) \geq \omega$?

Theorem 2. *Given $\alpha \in (\frac{n-1}{n}, 1)$,* MAX-WEIGHT-$\alpha$-SMCT *is NP-complete.*

Proof. It can easily be verified that MAX-WEIGHT-α-SMCT is in NP. Let $\alpha \in \left(\frac{n-1}{n}, 1\right)$. We define a polynomial-time reduction from SMT to α-SMCT.

Given an instance I of SMT, let $P_i = \{j_1^i, \ldots, j_n^i\}$ denote the preference list of i, which potentially includes ties. Let PG be an instance of α-SMCT with the same underlying graph as I. For each $(i,j) \in L \times R$, we define the cardinal preferences as $v_{ij} = -c_{ij} + n + 1$ and $w_{ji} = -c_{ji} + n + 1$. Note that as the costs are in the range $[1, n]$, so are the valuations by construction.

It trivially holds that M is a matching for PG if and only if M is a matching for I. We will prove that a matching M is α-stable for PG if and only if M is stable for I by showing that a pair (i,j) is α-blocking for PG if and only if it is blocking for I. Note that one direction holds by construction: take (i,j) to be α-blocking for PG, i.e. $v_{iM(i)} < \alpha v_{ij} < v_{ij}$ and $w_{jM(j)} < \alpha w_{ji} < w_{ji}$. This can only be the case if $c_{ij} < c_{iM(i)}$ and $c_{ji} < c_{jM(j)}$, so (i,j) is blocking for I. For the other direction, suppose that (i,j) is blocking for I, i.e. $c_{ij} < c_{iM(i)}$ and $c_{ji} < c_{jM(j)}$. This implies that $M(i)$ and j belong to different tie groups in i's preference list, and i and $M(j)$ belong to different tie groups in j's. We want to show that $v_{iM(i)} < \alpha v_{ij}$ and $w_{jM(j)} < \alpha w_{ji}$, thus it is enough to show that the valuation of an element in a tie group times α is greater than the valuation of any element in a following tie group. The most restrictive case happens when each tie group contains exactly one agent, in which case the largest ratio arises between the first and second agent in the preference list. More formally, let $2 \leq x \leq n$, then $\frac{v_{iM(i)}}{v_{ij}} \leq \frac{x-1}{x} \leq \frac{n-1}{n} < \alpha$, which indeed gives $v_{iM(i)} < \alpha v_{ij}$. The case of $w_{jM(j)} < \alpha w_{ji}$ follows analogously.

Take $K \in (1, 2n(n+1))$, and set $\omega = 2n(n+1) - K$. We now show that there exists an α-stable matching M for PG with $SW(M) \geq \omega$ if and only if there exists a stable matching M' for I with $w(M') \leq K$. Suppose that M is an α-stable matching for PG with $SW(M) \geq \omega$. It holds that

$$\begin{aligned} SW(M) &= \sum_{(i,j) \in M} (v_{ij} + w_{ji}) = \sum_{(i,j) \in M} (-c_{ij} - c_{ji} + 2n + 2) \\ &= -w(M) + 2n(n+1), \end{aligned}$$

where the last equality is due to M being perfect (since preference orders are complete). Thus, $w(M) \leq 2n(n+1) - \omega = K$. In particular, M is a stable matching for I with $w(M) \leq K$. Analogously, suppose that M is a stable matching for I with $w(M) \leq K$. It holds that

$$\begin{aligned} w(M) &= \sum_{(i,j) \in M} (c_{ij} + c_{ji}) = \sum_{(i,j) \in M} (-v_{ij} - w_{ji} + 2n + 2) \\ &= -SW(M) + 2n(n+1), \end{aligned}$$

and so $SW(M) \geq 2n(n+1) - K = \omega$. Thus M is an α-stable matching for PG with $SW(M) \geq \omega$. □

Note that we prove that MAX-WEIGHT-α-SMCT is NP-complete for $\alpha \in (\frac{n-1}{n}, 1)$. As mentioned above, the problem is solvable in polynomial time when $\alpha = 1$, and when $\alpha \leq \frac{\mu}{\mu+1}$ by Theorem 1. This reveals an interesting threshold phenomenon. Note that μ is determined by the valuations defined in the reduction and can be as small as $1/n$; the proof can be adapted to larger values of $\mu < 1$ by altering the way the valuations are defined. Clearly, Theorem 2 implies NP-hardness for more general optimization versions such as α-SMCTI.

6 Conclusion and Future Work

We study the interplay between stability and social welfare in two-sided matching markets with cardinal preferences, ties, and incomplete preference lists. By adopting the relaxed notion of α-stability, we provide a complete characterization of the stability–efficiency tradeoff under asymmetric valuations and show that relaxing stability can substantially improve achievable efficiency guarantees. We derive a polynomial-time algorithm that computes an α-stable matching attaining the best possible efficiency guarantee. Finally, our hardness results show that computing optimal α-stable matchings becomes intractable even under slight relaxations of stability, i.e., for $\alpha < 1$.

Our work leaves several interesting questions for future research. A natural direction is to further refine the complexity landscape of α-stable matchings. In particular, it would be interesting to determine whether computing an optimal α-stable matching is NP-hard for all $\alpha \in \left(\mu/\mu+1, 1\right)$, and to understand the approximability of optimal α-stable matchings within this range. Depending on these hardness results, a natural next step is to leverage predictions to guide the computation of near-optimal α-stable matchings; notably, our algorithm already draws inspiration from this paradigm. Other promising directions include extending the framework to many-to-one matching markets and exploring hybrid notions of stability that incorporate different behavioral assumptions.

References

[Anshelevich *et al.*, 2013] Elliot Anshelevich, Sanmay Das, and Yonatan Naamad. Anarchy, stability, and utopia: creating better matchings. *Autonomous Agents and Multi-Agent Systems*, 26(1):120–140, 2013.

[Budish, 2011] Eric Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.

[Caragiannis *et al.*, 2021] Ioannis Caragiannis, Aris Filos-Ratsikas, Panagiotis Kanellopoulos, and Rohit Vaish. Stable fractional matchings. *Artificial Intelligence*, 295:103416, 2021.

[Chen *et al.*, 2021] Jiehua Chen, Piotr Skowron, and Manuel Sorge. Matchings under preferences: Strength of stability and tradeoffs. *ACM Transactions on Economics and Computation*, 9(4):1–55, 2021.

[Chen *et al.*, 2025] Li Chen, Rasmus Kyng, Yang Liu, Richard Peng, Maximilian Probst Gutenberg, and Sushant Sachdeva. Maximum flow and minimum-cost flow in almost-linear time. *Journal of the ACM*, 72(3), 2025.

[Colini-Baldeschi *et al.*, 2024] Riccardo Colini-Baldeschi, Sophie Klumper, Guido Schäfer, and Artem Tsikiridis. To trust or not to trust: Assignment mechanisms with predictions in the private graph model. In *Proceedings of the 25th ACM Conference on Economics and Computation (EC)*, pages 1134–1154. ACM, 2024.

[Daskalakis *et al.*, 2009] Constantinos Daskalakis, Aranyak Mehta, and Christos Papadimitriou. A note on approximate Nash equilibria. *Theoretical Computer Science*, 410(17):1581–1588, 2009.

[Echenique *et al.*, 2025] Federico Echenique, Joseph Root, and Fedor Sandomirskiy. Stable matching as transport: a welfarist perspective on market design. *arXiv preprint arXiv:2402.13378*, 2025.

[Erdil and Ergin, 2017] Aytek Erdil and Haluk Ergin. Two-sided matching with indifferences. *Journal of Economic Theory*, 171:268–292, 2017.

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

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

[Gusfield and Irving, 1989] David Gusfield and Robert W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989.

[Haas, 2021] Christian Haas. Two-sided matching with indifferences: Using heuristics to improve properties of stable matchings. *Computational Economics*, 57:1115–1148, 2021.

[Iwama *et al.*, 1999] Kazuo Iwama, David Manlove, Shuichi Miyazaki, and Yasufumi Morita. Stable marriage with ties and incomplete lists. In *Proceedings the 26th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 1644 of *Lecture Notes in Computer Science*, pages 443–452. Springer, 1999.

[Knuth, 1997] Donald E. Knuth. *Stable Marriage and Its Relation to Other Combinatorial Problems: An Introduction to the Mathematical Analysis of Algorithms*, volume 10. American Mathematical Society, 1997.

[Lin *et al.*, 2024] Shiyun Lin, Simon Mauras, Nadav Merlis, and Vianney Perchet. Stable matching with ties: Approximation ratios and learning. *arXiv preprint arXiv:2411.03270*, 2024.

[Manlove *et al.*, 2002] David F. Manlove, Robert W. Irving, Kazuo Iwama, Shuichi Miyazaki, and Yasufumi Morita. Hard variants of stable marriage. *Theoretical Computer Science*, 276(1-2):261–279, 2002.

[Pini *et al.*, 2013] Maria Silvia Pini, Francesca Rossi, K. Brent Venable, and Toby Walsh. Stability, optimality and manipulation in matching problems with weighted preferences. *Algorithms*, 6(4):782–804, 2013.

[Roth and Sotomayor, 1990] Alvin E. Roth and Marilda A. O. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, volume 18. Cambridge University Press, 1990.

[Rothblum, 1992] Uriel G. Rothblum. Characterization of stable matchings as extreme points of a polytope. *Mathematical Programming*, 54:57–67, 1992.

[Troyan *et al.*, 2020] Peter Troyan, David Delacrétaz, and Andrew Kloosterman. Essentially stable matchings. *Games and Economic Behavior*, 120:370–390, 2020.

[Vande Vate, 1989] John H. Vande Vate. Linear programming brings marital bliss. *Operations Research Letters*, 8(3):147–153, 1989.