

A Simple 1.5-approximation Algorithm for a Wide Range of Maximum Size Stable Matching Problems

Gergely Csáji

HUN-REN Centre for Economic and Regional Studies

Budapest, Hungary

csaji.gergely@krtk.hu

ELTE Eötvös Loránd University

Budapest, Hungary

ABSTRACT

We give a simple approximation algorithm for a common generalization of many previously studied extensions of the maximum size stable matching problem with ties. These generalizations include the existence of critical vertices in the graph, amongst whom we must match as much as possible, free edges, that cannot be blocking edges and Δ -stabilities, which mean that for an edge to block, the improvement should be large enough on one or both sides. We also introduce other notions to generalize these even further, which allows our framework to capture many existing and future applications. We show that the edge duplicating technique allows us to treat these different types of generalizations simultaneously, while also making the algorithm, the proofs and the analysis much simpler and shorter than in previous approaches. In particular, we answer an open question by Askalidis et al. [1] about the existence of a $\frac{3}{2}$ -approximation algorithm for the MAX-SMTI problem with free edges. This demonstrates that this technique can grasp the underlying essence of these problems quite well and have the potential to be able to solve many future applications.

KEYWORDS

Matching; Approximation; Stable Matching; max-SMTI; Markets with preferences

ACM Reference Format:

Gergely Csáji. 2024. A Simple 1.5-approximation Algorithm for a Wide Range of Maximum Size Stable Matching Problems. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 7 pages.

1 INTRODUCTION

Preference based matching markets is an extensively studied topic in computer science, mathematics and economics literature. The intensive study of the area started back in 1962 after the seminal paper of Gale and Shapley [7]. They defined the model for the stable marriage problem and showed that a stable matching always exists and can be found in linear time. Since then, countless applications and related models have been studied, see Manlove [17] for

an overview. Applications include Resident allocation, University admissions, Kidney exchanges, job markets and much more.

1.1 Related work

In the stable marriage problem, the goal is to match two classes of agents in a way such that there are no blocking edges, i.e. pairs of agents who mutually consider each other better than their partner. The *stable marriage problem with ties and incomplete lists* (SMTI) was first studied by Iwama et al. [10], who showed the NP-hardness of MAX-SMTI, which is the problem of finding a maximum size stable matching in an SMTI instance. Since then, various algorithms have been proposed to improve the approximation ratio, e.g. Iwama et al. [11, 12] and Király [13], and the current best ratio is $\frac{3}{2}$ by a polynomial-time algorithm of McDermid [18], where the same ratio is attained by linear-time algorithms of Pauluch [20, 21] and Király [14, 15]. This $\frac{3}{2}$ -approximation has been extended to critical relaxed stable matchings by a very recent paper of Nasre et al. [19], which is a generalization of MAX-SMTI, where there is a set of critical agents, amongst whom we must match as much as possible and stability is generalized in an appropriate way. Critical relaxed stability has also been studied in Krishnaa et al. [16], where they showed that a critical relaxed stable matching always exists, but finding a maximum size such matching is NP hard to approximate within $\frac{21}{19} - \epsilon$, even with only strict preferences.

The approximation algorithm for MAX-SMTI has also been extended to some cardinal and matroidal generalizations as well by Csáji et al. [4]. This include for example common quotas and Δ -min-stability, where there is a number Δ , and blocking is relaxed in a way such that only those edges can block, where both participating agents improve by at least Δ . The matroidal generalization of stable matching was introduced in Fleiner [6].

As for the inapproximability of MAX-SMTI, Halldórsson et al. [8] showed that it is NP-hard to approximate it within some constant factor. Later, inapproximability results have been improved, especially assuming stronger complexity theoretic conjectures. Yanagisawa [22] and Halldórsson et al. [9] showed that assuming the Unique Games Conjecture, there is no $\frac{4}{3} - \epsilon$ -approximation for any $\epsilon > 0$. By a recent work by Dudycz et al. [5] it follows that assuming the Small Set Expansion Hypothesis or the strong-UGC, there cannot even be a $\frac{3}{2} - \epsilon$ -approximation algorithm for MAX-SMTI.

Stable matchings with free edges - that is, edges that can be included in a matching, but cannot block - have also been studied in many papers. Askalidis et al. [1] proved that with strict preferences and free edges, finding a maximum size stable matching can be $\frac{3}{2}$ -approximated in polynomial time. However, on the negative side



This work is licensed under a Creative Commons Attribution International 4.0 License.

they showed that the problem does not admit a polynomial time $\frac{21}{19} - \epsilon$ approximation algorithm unless $P = NP$ and assuming the Unique Games Conjecture, it does not even admit a $\frac{3}{2} - \epsilon$ approximation. These negative results also hold with respect to strict preferences. They posed as an open question whether this $\frac{3}{2}$ -approximation can be extended to weak preferences. In this paper we answer this question positively. Cechlárová and Fleiner [2] investigated the stable roommates problem with free edges and proved NP-hardness even in very restricted settings.

1.2 Our Contributions

In this paper we extend the $\frac{3}{2}$ -approximation algorithm for MAX-SMTI for a common generalization of many previously introduced and studied concepts. In particular, we answer an open question posed by Askalidis et al. [1] about the existence of a $\frac{3}{2}$ -approximation algorithm for the MAX-SMTI problem with free edges. We utilize and master a recently introduced edge duplicating technique by Yokoi [23] and Csáji et al. [4] and show that with the help of this technique, we are able to solve much harder and more general related problems in a lot simpler and more elegant way. We give a simple approximation algorithm for a framework that includes free edges, critical vertices, critical edges, many cardinal stability notions and more.

Although both our algorithm and our analysis are quite simple and elegant, we emphasize that our method can solve many problems that previously took whole papers in highly rated journals and conferences to tackle and also answers a decade long open question.

2 PRELIMINARIES

We investigate matching markets, where the set of agents with the possible set of contracts is given by a bipartite graph $G = (U, W; E)$, where vertices correspond to agents and edges correspond to possible, mutually acceptable contracts. For each agent $v \in U \cup W$, let $E(v)$ denote the edges that are incident to v . We assume that for each agent v , there is a *preference function* $p_v : E(v) \rightarrow \mathbb{R}_{\geq 0}$, which defines a weak ranking over the incident edges of v . We emphasize that we assume that the agents (vertices) rank their incident edges instead of the adjacent agents on the other side, because we allow parallel edges in our model, representing multiple types of contracts between two given agents. We also assume that $p_v(\emptyset) \leq 0$, which denotes that an agent always weakly prefers to be matched to any acceptable partners rather than being unmatched.

We say that an edge set $M \subset E$ is a *matching*, if $|M \cap E(v)| \leq 1$ for each $v \in U \cup W$. For a vertex v , let $M(v)$ denote the edge incident to v in M , if there is any, otherwise $M(v)$ is \emptyset . We say that an edge $e = (u, w) \notin M$ *blocks* a matching M , if $p_w(e) > p_w(M(w))$ and $p_u(e) > p_u(M(u))$ holds, that is, both agents strictly prefer each other to their partners in M . A matching M is called *weakly stable* or just *stable*, if there is no blocking edge to M . The problem of finding a maximum size weakly stable matching in a bipartite graph with weak preferences, called MAX-SMTI is a well studied problem that is NP-hard even to approximate within some constant, but admits a linear time $\frac{3}{2}$ -approximation.

We continue by introducing the generalizations that we consider. We start with Δ -min and Δ -max stability, which were introduced

in Csáji et al. [4] (a similar notion to Δ -min stability has also been defined in Chen et al. [3]). The motivation behind these notions is that in many applications, blocking pairs may only arise if one or both member of the pair improves by a significant enough amount.

Let $\Delta > 0$ be a constant. A matching M is *Δ -min-stable*, if there is no blocking edge $e = (u, w)$ with respect to M , such that $\min\{p_u(e) - p_u(M(u)), p_w(e) - p_w(M(w))\} \geq \Delta$, that is, both agents improve by at least Δ . The motivation behind this is clear: if a blocking does not result in a significant improvement, it may not even worth the effort.

A matching M is *Δ -max-stable*, if there is no blocking edge $e = (u, w)$ with respect to M , such that $\max\{p_u(e) - p_u(M(u)), p_w(e) - p_w(M(w))\} \geq \Delta$, that is, one of the improvements is at least Δ . Note here, that the edge e still must be a blocking edge, so both $p_u(e) - p_u(M(u)) > 0$ and $p_w(e) - p_w(M(w)) > 0$ are assumed. The motivation behind this notion is that in many cases, it may be enough that one agent can improve by a sufficiently large amount to initiate a blocking contract, and the other agent only has to accept it, so for him even a smaller improvement is enough. By choosing Δ to be small enough, it is easy to see that both of these stability notions strictly generalize weak stability.

Now we define critical matchings and relaxed stability, as they were defined in a very recent work of Krishnaa et al. [16]. For this, let us suppose that there is a set $C \subset U \cup W$ of *critical* vertices. Criticality means that we want to match as many of them as possible in any matching. We say that a matching M is *critical*, if there is no other matching M' , such that M' covers strictly more vertices from C , than M . Of course, there may not be a matching that is both critical and weakly stable. Hence, we need a weaker notion of stability, called *relaxed stability*. A matching M is *relaxed stable*, if there is no blocking pair $e = (u, w)$ to M , such that $M(u)$ and $M(w)$ are not critical vertices. The motivation behind this definition is that since we want to maximize the number of matched critical agents, we do not allow blocking pairs, where one of the participating agents would leave a critical agent unmatched. We will slightly strengthen this notion of stability for the following reason: this does not allow a pair consisting of an unmatched critical agent and another agent whose partner is critical to block. However, after letting the agents of such a blocking pair to be matched together, the number of critical agents covered remains the same, (so maximal), hence such a pair indeed should be able to block. We will show that this strengthening of stability is still enough to guarantee existence.

Now we generalize the concept of criticality and relaxed stability in a natural way. Suppose that there is a set E_c of critical edges too. This represents that it may be also important that we match the critical agents with specific types of contracts, or to certain other agents that are good enough for them, or more compatible with them in some way. For a matching M , let $M_c = M \cap E_c$ denote the critical edges of M . We say that a matching M is *critical*, if there is no matching N , such that N_c covers strictly more critical vertices than M_c . A matching M is *critical relaxed stable*, if it is critical and there is no blocking edge $e = (u, w)$, such that $M \setminus \{(u, M(u)), (M(w), w)\} \cup \{e\}$ is still critical. The special case of only critical vertices can be obtained by choosing $E = E_c$.

Finally we introduce the notion of free agents and free edges. In Askalidis et al. [1], free edges were used to model stable matching instances, where there are friendship and business relations which

restrict certain edges from blocking. Let us suppose there is a set $F \subset E$ of free edges and a set $F_A \subset U \cup W$ of free agents. Free agents cannot participate in any blocking edge, while free edges cannot be blocking edges. Hence, a matching M is *stable* in this instance, if for each blocking edge (u, w) it holds that $|\{u, w\} \cap F_A| \geq 1$ or $(u, w) \in F$. Clearly, we can model free agents with free edges only: just make each edge adjacent to a free agent a free edge.

Next we define a nice common generalization of Δ -min/max stability and the notion of free edges. Here we suppose that for each edge $e = (u, w)$, there are two numbers $\gamma_e^u > 0$ and $\gamma_e^w > 0$ given. We say that a matching is γ -*min stable*, if there is no blocking edge $e = (u, w)$ such that $p_u(e) - p_u(M(u)) \geq \gamma_e^u$ and $p_w(e) - p_w(M(w)) \geq \gamma_e^w$ holds. Similarly, a matching is γ -*max stable*, if there is no blocking edge $e = (u, w)$ such that $p_u(e) - p_u(M(u)) \geq \gamma_e^u$ or $p_w(e) - p_w(M(w)) \geq \gamma_e^w$ holds. The case of Δ -stabilities with a set $F \subset E$ of free edges corresponds to the very special case, when $\gamma_e^v = \infty$, if $e \in F$ and $\gamma_e^v = \Delta$ otherwise. Hence, in our model, we allow very different types of conditions for each edge to block, independently from each other, which can incorporate many other special properties of certain applications. For example, if one thinks about job markets, and assumes that the underlying preferences are in correspondence with the salaries of the positions, there may be many other aspects of a workplace that make it desirable. Hence, for each agent and each different company and position, the increase in the salary that would make a job offer good enough for the applicant to switch, might be different. Similarly, depending on a company's existing employees and a new agent, it may depend on the specific skills of a new agent that are relevant for the company, how large of a salary within a contract would be worth it for the company to switch from an existing employee to the new one in a blocking pair.

Putting it all together, we define a general model, which incorporates all of the previously discussed ones. Suppose we are given a set $C \subset U \cup W$ of critical vertices, a set $E_c \subset E$ of critical edges and γ_e^v values for each pair $(e, v) \in E \times (U \cup W)$ such that $v \in e$. We say that a matching M is γ -*min stable*, if M is critical and there is no blocking edge $e = (u, w)$, such that $M \setminus \{(u, M(u)), (M(w), w)\} \cup \{e\}$ is still critical, and $\min\{p_u(e) - p_u(M(u)) - \gamma_e^u, p_w(e) - p_w(M(w)) - \gamma_e^w\} \geq 0$ both hold. Otherwise, we call such a blocking edge a γ -*min blocking edge*. Clearly, this generalizes all three concepts: Δ -min stability, free edges and critical vertices. Similarly we can define γ -*max stable* matchings by replacing $\min\{p_u(e) - p_u(M(u)) - \gamma_e^u, p_w(e) - p_w(M(w)) - \gamma_e^w\} \geq 0$ with $\max\{p_u(e) - p_u(M(u)) - \gamma_e^u, p_w(e) - p_w(M(w)) - \gamma_e^w\} \geq 0$.

In order to be able to solve both problems with the same algorithm, let us go even one step further and consider a common generalization of the above two problems that we describe now. Instead of one γ_e^u , we have values $0 < \gamma_e^u < \delta_e^u$ for each vertex-edge pair. We say that an edge $e = (u, w)$ γ -*blocks* a matching M , if $M \setminus \{(u, M(u)), (M(w), w)\} \cup \{e\}$ is still critical, and either $\min\{p_u(e) - p_u(M(u)) - \gamma_e^u, p_w(e) - p_w(M(w)) - \delta_e^w\} \geq 0$ or $\min\{p_u(e) - p_u(M(u)) - \delta_e^u, p_w(e) - p_w(M(w)) - \gamma_e^w\} \geq 0$ holds. We say that a matching M is γ -*stable*, if M is critical and there is no edge that γ -blocks M .

If the γ_e^v values are sufficiently small, then this corresponds to γ -max stability, and when the γ_e^v and δ_e^v values are sufficiently close, then it corresponds to γ -min stability.

Now we define the optimization problem we investigate, called **MAXIMUM γ -STABLE MATCHING WITH TIES, INCOMPLETE PREFERENCES AND CRITICAL EDGES** abbreviated as **MAX- γ -SMTI** for short.

MAX- γ -SMTI

Input: A bipartite graph $G = (U, W; E)$, $p_v()$ preference valuations for each $v \in U \cup W$, a set $C \subset U \cup W$ of critical vertices, a set $E_c \subset E$ of critical edges, numbers $0 < \gamma_e^v < \delta_e^v$ for each pair $(e, v) \in E \times (U \cup W)$ such that $v \in e$.

Output: A maximum size γ -stable matching M

The main result of the paper is a simple $\frac{3}{2}$ -approximation algorithm for **MAX- γ -SMTI**.

3 THE ALGORITHM

We start by describing a high level view of the algorithm. Let I be an instance of **MAX- γ -SMTI**.

1. Create an instance I' of the stable marriage problem with strict preferences by making parallel copies of each edge and create strict preferences over the created edges.

2. Run the Gale-Shapley algorithm to obtain a stable matching M' in the new instance I'

3. Take the projection M of M' to I by taking an edge e inside M , whenever one of the parallel copies of e was inside M' .

The intuition behind this algorithm is the following. Our goal is a stable matching that is as large as possible. Hence, to help agents not to remain alone, we create copies for each edge, that allow rejected and unmatched agents multiple new chances to propose again to the same partner with a new, stronger contract that is more favorable to the other agent and can be good enough for her to reject another partner who has been considered better than the proposing agent before.

3.1 The case of $E_c = E$

Let us start with the special case when each edge is critical, so only the number of covered critical vertices matters. Also, this means that $M_c = M$ for any matching M . The more general case will be solved by the same simple ideas, however, the many required types of copies make it more difficult to follow. We describe the algorithm in more detail as follows. Let I be an instance of **MAX- γ -SMTI**. Let $|C \cap U| = s$ and $|C \cap W| = t$. For each edge $e = (u, w) \in E$ we create parallel edges as follows

- We create copies $a(e)$, $b_0(e)$, $b_1(e)$ and $c(e)$,
- If $w \in C \cap W$ is critical we create copies $x_1(e), \dots, x_t(e)$,
- If $u \in C \cap U$ is critical we create copies $z_1(e), \dots, z_s(e)$.

Then, we create strict preferences as follows. For a vertex $u \in U$, we rank the copies according to the rule

$$x_1 > \dots > x_t > a \geq b_0 \geq b_1 > c > z_s > \dots > z_1.$$

For a vertex $w \in W$, we rank the copies according to the rule

$$z_1 > \dots > z_s > c \geq b_1 \geq b_0 > a > x_t > \dots > x_1.$$

For any non- b_i copy, we rank the edges of the same copy according to the preference functions $p_v()$ by breaking the ties arbitrarily. Here, $\alpha > \beta$ denotes that for any two edges e, f , the copy $\alpha(e)$ is ranked higher than the copy $\beta(f)$.

The notation $a \geq^v b_0 \geq^{\delta-v} b_1$ means that for any edge e , $a(e) >_v b_0(e) >_v b_1(e)$, but $b_0(f) >_v a(e)$, if and only if $p_v(f) \geq p_v(e) + \gamma_f^v$ and $b_1(f) >_v a(e)$ if and only if $p_v(f) \geq p_v(e) + \delta_f^v$. This can be obtained in the following way. For each vertex $u \in U$ rank the a copies according to $p_v()$ by breaking the ties arbitrarily. Then, we can insert the b_0 copy of each edge f , such that $b_0(f) >_u a(e)$, if and only if $p_u(f) \geq p_u(e) + \gamma_f^u$. Similarly we can insert the b_1 copies such that $b_1(f) >_u a(e)$ if and only if $p_u(f) \geq p_u(e) + \delta_f^u$. For the W side, we do it similarly, according to their ranking. Note that this definition allows the b_i copies to be ranked in a different order than $p_v()$, which is surprising, but necessary. If there are still some ties remaining, we break them arbitrarily.

We remark another way to create the ranking over the a, b_0, b_1, c copies. For a vertex $u \in U$ and $e = (u, w)$ let $p_u(a(e)) = p_u(e)$, $p_u(b_0(e)) = p_u(e) - \gamma_e^u$ and $p_u(b_1(e)) = p_u(e) - \delta_e^u$. Then, for each vertex, these new (possibly negative) preference number define a weak order, in which we break ties in a way such that with the same value the b_1 copies are best, the b_0 copies are second and the a copies are last. We can do it similarly for the W side, with the tiebreaking b_0 first, b_1 second, c best.

We illustrate the creation of the extended preferences on a concrete example with the second method. Let $s = t = 1$. Let $u \in U$ be an agent with three adjacent edges e, f, g . Let $p_u(e) = 1$, $p_u(f) = 3$ and $p_u(g) = 4$. Also, let $\gamma_e^u = 1$, $\gamma_f^u = 2$, $\gamma_g^u = 2$ and $\delta_e^u = 2$, $\delta_f^u = 3$, $\delta_g^u = 6$. Then, $p_u(a(g)) = 4$, $p_u(a(f)) = 3$, $p_u(b_0(g)) = 2$, $p_u(b_0(f)) = p_u(a(e)) = 1$, $p_u(b_1(f)) = p_u(b_0(e)) = 0$, $p_u(b_1(e)) = -1$, $p_u(b_1(g)) = -2$. Then, the ranking between the copies is the following:

$$x_1(g) >_u x_1(f) >_u x_1(e) >_u a(g) >_u a(f) >_u b_0(g) >_u b_0(f) >_u a(e) >_u b_1(f) >_u b_0(e) >_u b_1(e) >_u b_1(g) >_u c(g) >_u c(f) >_u c(e) >_u z_1(g) >_u z_1(f) >_u z_1(e).$$

Notice that here $b_1(f) >_u b_1(e) >_u b_1(g)$, so for the b_1 copy, the edges are not ranked in their original order.

THEOREM 3.1. *MAX-CY-SMTI can be $\frac{3}{2}$ -approximated in $O((s+t)|E|)$ time, if $E = E_c$.*

PROOF. We prove the theorem in three simple claims. Let M denote the output of the algorithm and M' be its preimage in the extended instance I' .

Claim. M is critical.

Proof: Suppose for contradiction that there is a matching N , such that N covers strictly more critical vertices than M . Then, it must hold that either N covers more critical vertices from U or more critical vertices from W than M .

Suppose that the first case holds. Then, there is a component P (a path) in $N \cup M$, such that there is an endpoint $u_1 \in U \cap C$ that is critical, but is only matched in N , and also more critical vertices are covered in $N \cap P$, than in $M \cap P$. Let $e_1 = (u_1, w_1) = N(u_1)$. As u_1 is critical, we know that the copies $z_1(u_1, w_1), \dots, z_s(u_1, w_1)$ exist. As u_1 is not matched in M' , the fact that $z_1(u_1, w_1)$ does not block M' implies that w_1 is matched in M' and with a z_1 -type edge. Let $z_1(u_2, w_1) = M'(w_1)$. This immediately implies that $u_2 \in U \cap C$, as $z_1(u_2, w_1)$ exists. By our assumption on the component (more critical vertices are covered in $N \cap P$, than in $M \cap P$), u_2 is matched in N to a vertex w_2 .

As any z_2 copy is better for u_2 in I' , but $z_2(u_2, w_2)$ exists and does not block, $z_j(u_3, w_2) \in M'$ for some $u_3 \in U$ and $j \leq 2$. Again, we obtain that u_3 is critical, so by our assumption on the component, u_3 is matched in N to a vertex w_3 . By iterating this argument, we get that there must be vertices $u_1, u_2, \dots, u_{s+1} \in U$ that are critical, which contradicts $|U \cap C| = s$.

The second case is analogous. ■

Claim. The output matching M by the algorithm is cy -stable.

Proof: Suppose for contradiction that there is a cy -blocking edge $e = (u, w)$ to M .

If u is unmatched, then $g = (u', w) = M(w)$ exists. As e cy -blocks and M is critical, we get that $u \in C$ if and only if $u' \in C$. Hence, the best copy of e and g for w has the same type, either $z_1(e)$ or $c(e)$. In both cases, we get that this copy of e blocks M' as these copies are both ranked according to $p_w()$, contradiction. If w is unmatched, $x_1(e)$ or $a(e)$ blocks M' for similar reasons, contradiction.

Hence, $M(u) = f = (u, w')$ and $M(w) = g = (u', w)$ is not \emptyset . Because e cy -blocks, $p_u(e) \geq p_u(f) + \gamma_e^u$ and $p_w(e) \geq p_w(g) + \delta_e^w$ or $p_u(e) \geq p_u(f) + \delta_e^u$ and $p_w(e) \geq p_w(g) + \gamma_e^w$. Also, as e cy -blocks, the vertices u', w' cannot be critical, so $x_1(f), \dots, x_t(f), z_1(g), \dots, z_s(g)$ do not exist. Hence, in the first case $b_0(e)$ and in the second case $b_1(e)$ blocks M' , contradiction. ■

Claim. For any cy -min stable matching N it holds that $|N| \leq \frac{3}{2}|M|$.

Proof: Suppose for contradiction that there is a cy -stable matching N such that $|N| > \frac{3}{2}|M|$. Then, there must be a path component in $N \cup M$ that is either a single N -edge, or consists of two N -edges and one M -edge. The first case is clearly impossible, as M is necessarily maximal.

Suppose the second case holds. Let $e = (u_1, w_2)$ be the edge of M and $f = (u_1, w_1), g = (u_2, w_2)$ be the edges of N .

First observe that as M and N are critical, e and $f \cup g$ cover the same number of critical vertices, so w_1, u_2 are not critical and $x_1(f), \dots, x_t(f), z_1(g), \dots, z_s(g)$ does not exist.

As $a(f)$ does not block, $b_0(e), b_1(e), a(e)$ or $x_i(e) \in M'$ for some i . As $c(g)$ does not block, $b_0(e), b_1(e), c(e)$ or $z_j(e) \in M'$ for some j . Hence, $b_0(e)$ or $b_1(e) \in M'$. If $b_0(e) \in M'$, then $p_{u_1}(e) \geq p_{u_1}(f) + \gamma_e^{u_1}$ and $p_{w_2}(e) \geq p_{w_2}(g) + \delta_e^{w_2}$ as M' is stable, so e cy -blocks N (combining with the fact that e and $f \cup g$ cover the same number of critical vertices), contradiction. The other case implies $p_{u_1}(e) \geq p_{u_1}(f) + \delta_e^{u_1}$ and $p_{w_2}(e) \geq p_{w_2}(g) + \gamma_e^{w_2}$, so e cy -blocks N , contradiction again. ■

The statement follows from these three claims. The running time of the algorithm is linear in the number of edges of the extended instance, as the Gale-Shapley algorithm is linear, so it has running time $O((s+t)|E|)$. □

3.2 General case

Next, we discuss the more general case with an arbitrary subset of critical edges.

We may assume without loss of generality that each critical edge E_c is adjacent to at least one critical vertex, otherwise it cannot cover any critical vertices, so we get an equivalent instance by making the edge non-critical.

Let us start by describing the extended instance I' . For each edge $e = (u, w) \in E$ we create parallel edges as follows:

- We create copies $a(e), b_0(e), b_1(e)$ and $c(e)$,
- If $w \in C \cap W$ is critical and $e \in E_c$ is also critical, we create copies $x_1(e), \dots, x_{t+7}(e)$
- If $u \in C \cap U$ is critical and $e \in E_c$ is also critical, we create copies $z_1(e), \dots, z_{s+7}(e)$
- Finally, if $e \in E_c$ is critical and both $u, w \in C$ are critical, we create a copies $y_0(e), y_1(e)$

Then, we create strict preferences as follows. Let $\delta' = \delta - \gamma$. For a vertex $u \in U$, we rank the copies according to the rule

$$x_1 \geq^Y x_2 \geq^{\delta'} x_3 > x_4 > \dots > x_{t+4} > z_{s+7} \geq^Y (y_0 \geq z_{s+6}) \geq^{\delta'} (y_1 \geq z_{s+5}) > z_{s+4} > \dots > z_4 > a \geq^Y (b_0 \geq z_3 \geq x_{t+5}) \geq^{\delta'} (b_1 \geq z_2 \geq x_{t+6}) > z_1 > x_{t+7} > c.$$

For a vertex $w \in W$, we rank the copies according to the rule

$$z_1 \geq^Y z_2 \geq^{\delta'} z_3 > z_4 > \dots > z_{s+4} > x_{t+7} \geq^Y (y_1 \geq x_{t+6}) \geq^{\delta'} (y_0 \geq x_{t+5}) > x_{t+4} > \dots > x_4 > c \geq^Y (b_1 \geq x_3 \geq z_{s+5}) \geq^{\delta'} (b_0 \geq x_2 \geq z_{s+6}) > x_1 > z_{s+7} > a.$$

Here, when we denote $(\alpha \geq \beta)$ for some types of copies, then it always means that we insert $\alpha(e)$ and $\beta(e)$ together with $\beta(e)$ being strictly after $\alpha(e)$ (so no other edge comes between them) and we keep them together, so in the end for any other edge, either it is worse than both of them or better than both of them. In the model where we assign preference values to the copies to create a new ranking, this corresponds to that these copies always get the same value, so they are tied, and then we keep them together in the tiebreaking in that order.

The only important thing is that, if we have some copies $\alpha \geq^Y (\beta_1 \geq \dots \geq \beta_k) \geq^{\delta'} (\omega_1 \geq \dots \geq \omega_l)$, then for any edge e , $\alpha(e) >_v \beta_1(e) >_v \dots >_v \beta_k(e) >_v \omega_1(e) >_v \dots >_v \omega_l(e)$ and that for any $i \in \{1, \dots, k\}$ and $j \in \{1, \dots, l\}$ it holds that $\beta_i(f) >_v \alpha(e)$, if and only if $p_v(f) \geq p_v(e) + \gamma_f^v$ and $\omega_j(f) >_v \alpha(e)$ if and only if $p_v(f) > p_v(e) + \delta_f^v$.

The edges from copies $x_1, x_4, x_5, \dots, x_{t+4}, x_{t+7}, z_1, z_4, z_5, \dots, z_{s+4}, z_{s+7}, a, c$ are ranked according to $p_v()$, the others are then inserted to the right place as before. In the preference value model, we do this by computing the corresponding values similarly as in the previous case (with each copy in a parenthesis getting the same value) and for tie breaking we use a suitable rule. For example for a vertex $u \in U$, amongst the copies $a, b_0, b_1, z_2, z_3, x_{t+5}, x_{t+6}$, we can use $b_1 >_u z_2 >_u x_{t+6} >_u b_0 >_u z_3 >_u x_{t+5} >_u a$. The ties between the same type of copies can be broken arbitrarily.

THEOREM 3.2. *MAX-CY-SMTI can be $\frac{3}{2}$ -approximated in $O((s+t)|E|)$ time.*

PROOF. We prove the theorem in three simple claims. Let M denote the output of the algorithm and M' be its preimage in the extended instance I' .

Claim. M is critical.

Proof: Suppose for contradiction that there is a matching N , such that N_c covers strictly more critical vertices than M_c . Then, it must hold that either N_c covers more critical vertices from U or more critical vertices from W than M_c .

Suppose that the first case holds. Then, there is a component (a path) in $N_c \cup M_c$, such that there is an endpoint $u_1 \in U \cap C$ that is

critical, but is only covered in N_c and also more critical vertices are covered in $N_c \cap P$, than in $M_c \cap P$. Let $e_1 = (u_1, w_1) = N_c(u_1)$. As u_1 is critical, we know that the copies $z_1(u_1, w_1), \dots, z_{s+7}(u_1, w_1)$ exist. As u_1 is not covered with a critical edge in M (as it is uncovered in M_c), it can only be matched with an a, b_j or c type copy. The fact that $z_4(u_1, w_1)$ does not block M' implies that w_1 is matched in M' and with a z_j -type edge with $j \leq 4$. Let $(u_2, w_1) = M(w_1)$. This immediately implies that $u_2 \in U \cap C$ and $(u_2, w_1) \in E_c$, as $z_j(u_2, w_1)$ exists. By our assumption on the component (more critical vertices are covered in $N_c \cap P$, than in $M_c \cap P$), u_2 is matched in N_c to a vertex w_2 .

As any z_5 copy is better for u_2 in I' than any z_j copy with $j \leq 4$, but $z_5(u_2, w_2)$ exists and does not block, $z_j(u_3, w_2) \in M'$ for some $u_3 \in U$ and $j \leq 5$. Again, we obtain that u_3 is critical and $(u_3, w_2) \in E_c$, so by our assumption on the component, u_3 is matched in N_c to a vertex w_3 . By iterating this argument, we get that there must be vertices $u_1, u_2, \dots, u_{s+1} \in U$ that are critical, which contradicts $|U \cap C| = s$.

The second case is analogous. ■

Claim. The output matching M by the algorithm is cy -stable.

Proof: Suppose for contradiction that there is a cy -blocking edge $e = (u, w)$ to M .

If u is unmatched, then $g = M(w) = (u', w)$ exists. As e cy -blocks and M is critical we get that $e \in E_c$ if and only if $g \in E_c$ and if yes, then $u' \in C$ if and only if $u \in C$. In particular, we get that the best copy of e and g in I' for w is the same. Also, this copy can be only c, x_{t+7} or z_1 and all these copies are ranked according to $p_w()$. Therefore, this copy of e blocks M' , as e cy -blocks M , contradiction. The case when w is unmatched is similar.

Hence, $M(u) = f = (u, w')$ and $M(w) = g = (u', w)$ is not \emptyset . Because e cy -blocks, $p_u(e) \geq p_u(f) + \gamma_e^u$ and $p_w(e) \geq p_w(g) + \delta_e^w$ or $p_u(e) \geq p_u(f) + \delta_e^u$ and $p_w(e) \geq p_w(g) + \gamma_e^w$.

If $f, g \notin E_c$, then $e \notin E_c$, as M is critical. Hence, only a, b_0, b_1, c copies exists of e, f, g . As e cy -blocks M , we get that $b_0(e)$ or $b_1(e)$ blocks M' , contradiction.

Suppose that $f, g \in E_c$. Then, $e \in E_c, u, w \in C$ and $u', w' \notin C$, because e cy -blocks. Hence, $x_i(f), z_j(g)$ do not exists for any i, j , but $y_0(e), y_1(e)$ exist. As e cy -blocks one of $y_0(e), y_1(e)$ blocks M' , contradiction.

Suppose that $f \in E_c, g \notin E_c$. In particular, only $a(g), b_0(g), b_1(g), c(g)$ exist of g . Then, $e \in E_c$ as e cy -blocks and $w' \in C$ if and only if $w \in C$. If $w, w' \in C$, then $x_2(e), x_3(e)$ exist and one of them blocks M' , contradiction. If $w, w' \notin C$, then $u \in C$ (by our assumption for any $e \in E_c$, at least one endpoint is critical) and $x_i(e), x_i(f)$ do not exists for any i and neither does $y_0(e), y_0(f), y_1(e), y_1(f)$. Hence, $z_{s+5}(e)$ or $z_{s+6}(e)$ blocks M' , contradiction.

The case $f \notin E_c, g \in E_c$ is analogous. ■

Claim. For any cy -min stable matching N it holds that $|N| \leq \frac{3}{2}|M|$.

Proof: Suppose for contradiction that there is a cy -stable matching N such that $|N| > \frac{3}{2}|M|$. Then, there must be a path component in $N \cup M$ that is either a single N -edge, or consist of two N -edges and one M -edge. The first case is clearly impossible, as M is necessarily maximal.

Suppose the second case holds. Let $e = (u_1, w_2)$ be the edge of M and $f = (u_1, w_1)$, $g = (u_2, w_2)$ be the edges of N .

First observe that as M and N are critical, e and $f \cup g$ cover the same number of critical vertices with critical edges. Hence, if e γ -blocks N , then it γ -blocks N .

Suppose that $f, g \notin E_c$. Then, $e \notin E_c$ and only a, b_0, b_1, c copies exists of all three edges. As $a(f)$ does not block M' , $b_0(e), b_1(e)$ or $a(e) \in M'$. As $c(g)$ does not block M' , $b_0(e), b_1(e)$ or $c(e) \in M'$. Hence, $b_0(e)$ or $b_1(e) \in M'$. Furthermore, if $b_0(e) \in M'$, then $p_{u_1}(e) \geq p_{u_1}(f) + \gamma_e^{u_1}$ and $p_{w_2}(e) \geq p_{w_2}(g) + \delta_e^{w_2}$, so e γ -blocks N (combining with the fact that e and $f \cup g$ cover the same number of critical vertices), contradiction. The other case implies $p_{u_1}(e) \geq p_{u_1}(f) + \delta_e^{u_1}$ and $p_{w_2}(e) \geq p_{w_2}(g) + \gamma_e^{w_2}$, so e γ -blocks N , contradiction again.

Next suppose that $f, g \in E_c$. Then, as M is critical, $e \in E_c$, $u_1, w_2 \in C$ and $u_2, w_1 \notin C$. Hence, $x_i(f), z_j(g)$ does not exist for any i, j but $y_0(e), y_1(e)$ exist and also $z_j(f), x_i(g)$ exists for all i, j . As $z_{s+7}(f)$ does not block and $x_{t+7}(g)$ does not block M' we get that $y_0(e) \in M'$ or $y_1(e) \in M'$ and in both cases, e γ -blocks N , contradiction.

Suppose that $f \in E_c, g \notin E_c$. Then, as M, N are critical, $e \in E_c$ and $w_1 \in C$ if and only if $w_2 \in C$. If $w_1, w_2 \in C$, then $x_1(f)$ exists. As $x_1(f)$ does not block and $c(g)$ does not block M' , $x_2(e) \in M'$ or $x_3(e) \in M'$. In both cases, e γ -blocks N , contradiction. If $w_1, w_2 \notin C$, then $u_1 \in C$ (each $e \in E_c$ has one critical endpoint by our assumption) and $y_0(e), y_1(e), x_i(e)$ does not exist for $i \in \{1, \dots, t+7\}$, but $z_j(e), z_j(f)$ does for $j \in \{1, \dots, s+7\}$. As neither $z_{s+7}(f)$ nor $c(g)$ blocks M' , we get that $z_{s+6}(e) \in M'$ or $z_{s+5}(e) \in M'$. In both cases, we get that e γ -blocks N , contradiction.

The case $f \notin E_c, g \in E_c$ is similar. ■

The statement follows from these three claims. The running time of the algorithm is linear in the edges of the extended instance, as the Gale-Shapley algorithm is linear, so it has running time $O((s+t)|E|)$. □

We state one remark about a straightforward extension of this algorithm.

Remark. If we can have k $(\gamma_e^v)_1 < (\gamma_e^v)_2 < \dots < (\gamma_e^v)_k$ values instead of the two γ_e^v, δ_e^v values, and let $e = (u, w)$ block if $p_u(e) \geq p_u(M(u)) + (\gamma_e^u)_i$ and $p_w(e) \geq p_w(M(w)) + (\gamma_e^w)_{k+1-i}$ for some $i \in \{1, \dots, k\}$, then this framework can also incorporate a notion of γ -sum stability. In γ -sum stability, we have one γ_e value for each edge, and (u, w) γ -sum blocks, if it blocks and the sum of the two improvements are at least γ_e . With $k = O(|U \cup W|^2)$ copies, we can have $(\gamma_e^v)_1 < (\gamma_e^v)_2 < \dots < (\gamma_e^v)_k$ values for each possible improvement c and $\gamma_e - c$ such that $(\gamma_e^v)_i = \gamma_e - (\gamma_e^v)_{k+1-i}$, so γ -sum stability becomes a special case. Furthermore, the same algorithm straightforwardly extends to this case, we only need to make k copies instead of 2 for the needed cases and define the strict ranking in I' according to the differences $(\gamma_e^v)_{i+1} - (\gamma_e^v)_i$.

4 CONCLUSION AND FUTURE WORK

In this paper we demonstrated the robust usefulness and generality of the edge duplicating technique and provided simple algorithms for many MAX-SMTI generalizations at the same time. A natural

direction to generalize the algorithm even further is to allow capacities or matroid constraints for the vertices. Our preliminary results show that our algorithm indeed extends to these cases. As there are probably even more cases, where this technique could provide a nice solution, there is a huge potential of applying it to other lesser known or yet to be introduced models.

ACKNOWLEDGMENTS

The author was supported by the Hungarian Scientific Research Fund, OTKA, Grant No. K143858, by the Momentum Grant of the Hungarian Academy of Sciences, grant number 2021-1/2021 and by the Doctoral Student Scholarship Program (number C2258525) of the Cooperative Doctoral Program of the Ministry of Innovation and Technology financed by the National Research, Development, and Innovation Fund.

REFERENCES

- [1] Georgios Askalidis, Nicole Immorlica, Augustine Kwanashie, David F Manlove, and Emmanouil Pountourakis. 2013. Socially stable matchings in the hospitals/residents problem. In *Algorithms and Data Structures: 13th International Symposium, WADS 2013, London, ON, Canada, August 12-14, 2013. Proceedings 13*. Springer, 95–96.
- [2] Katarina Cechlárová and Tamás Fleiner. 2009. *Stable roommates with free edges*. Technical Report. Technical Report 2009-01, Egerváry Research Group on Combinatorial Optimization.
- [3] Jiehua Chen, Piotr Skowron, and Manuel Sorge. 2021. Matchings under preferences: Strength of stability and tradeoffs. *ACM Transactions on Economics and Computation* 9, 4 (2021), 1–55.
- [4] Gergely Csáji, Tamás Király, and Yu Yokoi. 2023. Approximation Algorithms for Matroidal and Cardinal Generalizations of Stable Matching. In *Symposium on Simplicity in Algorithms (SOSA)*. SIAM, 103–113.
- [5] Szymon Dudycz, Pasin Manurangsi, and Jan Marcinkowski. 2022. Tight Inapproximability of Minimum Maximal Matching on Bipartite Graphs and Related Problems. In *Approximation and Online Algorithms: 19th International Workshop, WAOA 2021, Lisbon, Portugal, September 6–10, 2021, Revised Selected Papers*. Springer, 48–64.
- [6] Tamás Fleiner. 2001. A matroid generalization of the stable matching polytope. In *Proc. 8th International Conference on Integer Programming and Combinatorial Optimization*. Springer, 105–114.
- [7] David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.
- [8] Magnús Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Yasufumi Morita. 2002. Inapproximability results on stable marriage problems. In *LATIN 2002: Theoretical Informatics: 5th Latin American Symposium Cancun, Mexico, April 3–6, 2002 Proceedings 5*. Springer, 554–568.
- [9] Magnús M Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. 2003. Improved approximation of the stable marriage problem. In *Algorithms-ESA 2003: 11th Annual European Symposium, Budapest, Hungary, September 16-19, 2003. Proceedings 11*. Springer, 266–277.
- [10] Kazuo Iwama, David Manlove, Shuichi Miyazaki, and Yasufumi Morita. 1999. Stable marriage with incomplete lists and ties. In *Proc. 26th International Colloquium on Automata, Languages, and Programming (ICALP 1999)*. Springer, 443–452.
- [11] Kazuo Iwama, Shuichi Miyazaki, and Naoya Yamauchi. 2007. A 1.875-approximation algorithm for the stable marriage problem. In *Proc. Eighteenth annual ACM-SIAM symposium on Discrete algorithms (SODA 2007)*. SIAM, Philadelphia, 288–297.
- [12] Kazuo Iwama, Shuichi Miyazaki, and Naoya Yamauchi. 2008. A $(2 - c \frac{1}{\sqrt{n}})$ -Approximation Algorithm for the Stable Marriage Problem. *Algorithmica* 51, 3 (2008), 342–356.
- [13] Zoltán Király. 2011. Better and simpler approximation algorithms for the stable marriage problem. *Algorithmica* 60, 1 (2011), 3–20.
- [14] Zoltán Király. 2012. Linear time local approximation algorithm for maximum stable marriage. In *Proc. Second International Workshop on Matching Under Preferences (MATCH-UP 2012)*. 99.
- [15] Zoltán Király. 2013. Linear time local approximation algorithm for maximum stable marriage. *Algorithms* 6, 3 (2013), 471–484.
- [16] Prem Krishnaa, Girija Limaye, Meghana Nasre, and Prajakta Nimbhorkar. 2023. Envy-freeness and relaxed stability: Hardness and approximation algorithms. *Journal of Combinatorial Optimization* 45, 1 (2023), 41.
- [17] David Manlove. 2013. *Algorithmics of matching under preferences*. Vol. 2. World Scientific.

- [18] Eric McDerimid. 2009. A $3/2$ -approximation algorithm for general stable marriage. In *Proc. 36th International Colloquium on Automata, Languages, and Programming (ICALP 2009)*. Springer, 689–700.
- [19] Meghana Nasre, Prajakta Nimbhorkar, and Keshav Ranjan. 2023. Critical Relaxed Stable Matchings with Two-Sided Ties. *arXiv preprint arXiv:2303.12325* (2023).
- [20] Katarzyna Paluch. 2011. Faster and simpler approximation of stable matchings. In *Proc. 9th International Workshop on Approximation and Online Algorithms (WAOA 2011)*. 176–187.
- [21] Katarzyna Paluch. 2014. Faster and simpler approximation of stable matchings. *Algorithms* 7, 2 (2014), 189–202.
- [22] Hiroki Yanagisawa. 2007. Approximation algorithms for stable marriage problems. *PhD thesis, Kyoto University, Graduate School of Informatics* (2007).
- [23] Yu Yokoi. 2021. An approximation algorithm for maximum stable matching with ties and constraints. *arXiv preprint arXiv:2107.03076* (2021).