Fairly Popular Matchings and Optimality

Telikepalli Kavitha
Tata Institute of Fundamental Research, Mumbai, India

Abstract

We consider a matching problem in a bipartite graph $G = (A \cup B, E)$ where vertices have strict preferences over their neighbors. A matching M is popular if for any matching N, the number of vertices that prefer M is at least the number that prefer N; thus M does not lose a head-to-head election against any matching where vertices are voters. It is easy to find popular matchings; however when there are edge costs, it is NP-hard to find (or even approximate) a min-cost popular matching. This hardness motivates relaxations of popularity.

Here we introduce *fairly popular* matchings. A fairly popular matching may lose elections but there is no good matching (wrt popularity) that defeats a fairly popular matching. In particular, any matching that defeats a fairly popular matching does not occur in the support of any popular mixed matching. We show that a min-cost fairly popular matching can be computed in polynomial time and the fairly popular matching polytope has a compact extended formulation.

We also show the following hardness result: given a matching M, it is NP-complete to decide if there exists a popular matching that defeats M. Interestingly, there exists a set K of at most m popular matchings in G (where $|E| = m$) such that if a matching is defeated by some popular matching in G then it has to be defeated by one of the matchings in K.

2012 ACM Subject Classification Theory of computation → Design and analysis of algorithms

Keywords and phrases Bipartite graphs, Stable matchings, Mixed matchings, Polytopes

Digital Object Identifier 10.4230/LIPIcs.STACS.2022.41

Funding *Telikepalli Kavitha*: Supported by the Department of Atomic Energy, Government of India, under project no. RTI4001.

Acknowledgements Thanks to the reviewers for their helpful comments and suggestions. Thanks to Yuri Faenza and Jaikumar Radhakrishnan for useful discussions.

1 Introduction

Our input is a bipartite graph $G = (A \cup B, E)$ on n vertices and m edges where every vertex has a strict ranking of its neighbors. Such a graph is also called a marriage instance and this is a very well-studied model in two-sided matching markets. A matching M is stable if no edge *blocks* it; edge (a, b) blocks M if both a and b prefer each other to their respective assignments in M. The existence of stable matchings in a marriage instance and the Gale-Shapley algorithm [14] to find one are classical results in algorithms.

Stable matchings are used in many real-world applications such as matching students to schools and colleges [1, 3] and medical residents to hospitals [5, 28]. Stability is a rather strict notion – all stable matchings match the same subset of vertices [15] and the size of a stable matching might be only half the size of a maximum matching. In several applications, the notion of stability can be relaxed to a less demanding notion for the sake of collective welfare.

Popularity is a meaningful relaxation of stability based on empowering *matchings* (instead of edges) to block other matchings. Any pair of matchings, say M and N, can be compared by holding an election between them where every vertex v either casts a vote for the matching in $\{M, N\}$ where it gets a better partner (and being unmatched is its worst choice) or v

© Telikepalli Kavitha;
licensed under Creative Commons License CC-BY 4.0
39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022).
Editors: Petra Berenbrink and Benjamin Monmege; Article No. 41; pp. 41:1–41:22
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

abstains from voting if it is indifferent between M and N. Let $\phi(M,N)$ (resp., $\phi(N,M)$) be the number of votes for M (resp., N). Matching N is *more popular* than matching M (equivalently, N *defeats* M) if $\phi(N,M) > \phi(M,N)$. Let $\Delta(M,N) = \phi(M,N) - \phi(N,M)$.

▶ **Definition 1.** *A matching M is popular if there is no matching more popular than M, i.e., $\Delta(M,N) \geq 0$ for all matchings N in G.*

Gärdenfors [16] introduced the notion of popularity in 1975 where he showed that every stable matching is popular. In fact, stable matchings are min-size popular matchings [18]. Hence relaxing stability to popularity allows larger matchings and more generally, matchings with lower cost (when every edge has a cost) to be feasible.

Several algorithmic and hardness results for popular matchings have been obtained during the last decade and we refer to [6] for a survey. We know efficient algorithms for only a few popular matching problems such as the max-size popular matching problem and the popular edge problem [7, 18, 21]. Many natural optimization problems in popular matchings such as the min-cost popular matching problem are NP-hard [10]; moreover, this problem is NP-hard to approximate to any multiplicative factor. Though relaxing stability to popularity promises matchings with improved optimality, finding these matchings is hard.

The extension complexity of the popular matching polytope of G is $2^{\Omega(m/\log m)}$ [9]. Thus formulating the convex hull of edge incidence vectors of matchings M that satisfy $\Delta(M,N) \geq 0$ for *all* matchings N is hard. This motivates relaxing popularity, i.e., let us waive some constraints $\Delta(M,N) \geq 0$. For what matchings N would it be justified to do so?

Suppose N is "very unpopular" – then N is not a viable alternative and it seems fair to not give N the power to block other matchings. Forbidding *very unpopular* matchings from blocking others is similar in spirit to legal assignments [8] (a relaxation of stable matchings) where only edges that belong to legal assignments are allowed to block matchings. Thus our goal is to come up with a filter that tests matchings for a natural relaxation of popularity and forbid the ones that fail our test to block matchings.

So we seek to identify a subset $\mathcal{S}$ of the set of all matchings in G such that:

(a) Every matching outside $\mathcal{S}$ fails our test that checks for "mild popularity".
(b) It is easy to optimize over matchings M that satisfy $\Delta(M,N) \geq 0$ for all $N \in \mathcal{S}$.
(c) For any matching $T \notin \mathcal{S}$, there is at least one matching $N \in \mathcal{S}$ such that $\Delta(T,N) < 0$.

▶ Remark 2. Note that property (c) is independent of property (a); the latter says every matching $T \notin \mathcal{S}$ has to fail our test of *mild popularity* while the former says any matching $T \notin \mathcal{S}$ has to be defeated by a matching in $\mathcal{S}$, so we will not have $\Delta(T,N) \geq 0$ for all $N \in \mathcal{S}$.

The unpopularity of a matching T is typically measured by its unpopularity factor [27], defined as $u(T) = \max_{N \neq T} \phi(N,T)/\phi(T,N)$. A matching T is popular if and only if $u(T) \leq 1$. Suppose we define a matching T to be very unpopular if $u(T) > k$ for some k. Is it easy to compute a min-cost matching M such that $\Delta(M,N) \geq 0$ for all matchings N with $u(N) \leq k$?

When $k = n-1$, it means that no Pareto optimal matching defeats M – observe that such a matching M has to be popular. So the above problem is NP-hard for $k = n-1$. We show this problem is coNP-hard for $k = 1$ (see Remark 9). Thus using unpopularity factor to come up with a test of mild popularity does not look very promising for tractability.

Our main result. Rather than unpopularity factor, we will use popular *mixed* matchings [26] to define a natural relaxation of popularity. A mixed matching Π is a probability distribution or a lottery over matchings, so $\Pi = \{(M_0,p_0),\ldots,(M_k,p_k)\}$ where $M_0,\ldots,M_k$ are matchings, $p_i > 0$ for all i, and $\sum_{i=0}^{k} p_i = 1$. The notion of popularity can be extended to mixed matchings; the mixed matching Π is popular if $\Delta(\Pi,N) = \sum_{i=0}^{k} p_i \cdot \Delta(M_i,N) \geq 0$ for all matchings N.

The matchings $M_0, \ldots, M_k$ are said to be in the support of $\Pi = \{(M_0, p_0), \ldots, (M_k, p_k)\}$. Let us call a matching M *supporting* if there exists a popular mixed matching Π whose support contains M. So every supporting matching participates in some popular lottery over matchings, thus the "supporting" property is a natural relaxation of popularity – we will use this property as our condition for mild popularity. We define *fairly popular* matchings now.

▶ **Definition 3.** *A matching M is fairly popular if $\Delta(M, N) \geq 0$ $\forall$ supporting matchings N.*

For any matching T that defeats a fairly popular matching M, it is the case that even with the help of other matchings, T cannot form a popular mixture. Thus it is natural to regard a *non-supporting* matching T as being "very unpopular". So we set the supporting property as our threshold for mild popularity – thus elections against non-supporting matchings will not be relevant. In other words, even if $\Delta(M, T) < 0$ for a non-supporting matching T, the matching M will continue to be feasible. Intriguingly, waiving the constraints $\Delta(M, T) \geq 0$ for non-supporting matchings T makes the resulting polytope easy to describe.

▶ **Theorem 4.** *Given a marriage instance $G = (A \cup B, E)$ with edge costs, a min-cost fairly popular matching can be computed in polynomial time. Furthermore, the convex hull of edge incidence vectors of fairly popular matchings has a compact extended formulation.*

Key to the above theorem is our characterization of supporting matchings (see Theorem 5). Any point $x \in \mathbb{R}^m_{\geq 0}$ such that $\sum_{e \in \delta(v)} x_e \leq 1$ for each vertex v is a *fractional* matching and x is equivalent to a mixed matching (Birkhoff-von Neumann theorem). A fractional matching x is popular if Π is a popular mixed matching, where Π is any mixed matching that corresponds to x (see [26]). An edge e is a *popular fractional* edge if there exists a popular fractional matching x with $x_e > 0$. Let $E_p \subseteq E$ be the set of popular fractional edges.

Let us call a vertex v *stable* if v is matched in any (equivalently, every [15]) stable matching in G. So unstable vertices are those left unmatched in every stable matching.

▶ **Theorem 5.** *Let $G = (A \cup B, E)$ be a marriage instance and let M be a matching in G. The following three statements are equivalent.*
1. *M is supporting, i.e., M occurs in the support of some popular mixed matching.*
2. *No popular mixed matching defeats M, i.e., $\Delta(\Pi, M) = 0$ $\forall$ popular mixed matchings Π.*
3. *M matches all stable vertices and $M \subseteq E_p$.*

▶ Remark 6. Theorem 5 implies that any matching that is *non-supporting* is defeated by some popular mixed matching and thus, by some supporting matching (since every popular mixed matching is a lottery over supporting matchings). So $\mathcal{S} = \{\text{supporting matchings}\}$ satisfies properties (a), (b), and (c) stated earlier. Thus every fairly popular matching is also supporting.

Observe that the set of popular matchings does not satisfy the property that any matching outside this set has to be defeated by at least one matching in this set. That is, it is not the case that every *unpopular* matching has to lose to one or more popular matchings. For example, consider the following instance where $A = \{a_0, a_1, a_2\}$ and $B = \{b_0, b_1\}$.

$a_0\colon\ b_0 \succ b_1$	$a_1\colon\ b_0 \succ b_1$	$a_2\colon\ b_1$
$b_0\colon\ a_0 \succ a_1$	$b_1\colon\ a_0 \succ a_1 \succ a_2$	

Here a_0 and b_0 are each other's top choice neighbors and a_0's second choice is b_1 and b_0's second choice is a_1 and so on. The above instance has only one popular matching $P = \{(a_0, b_0), (a_1, b_1)\}$. The matching $M = \{(a_0, b_1), (a_1, b_0)\}$ is not popular since the matching $N = \{(a_0, b_0), (a_2, b_1)\}$ is more popular than M; the vertices a_0, b_0, a_2 prefer N while a_1, b_1 prefer M. Observe that the popular matching P is *not* more popular than M.

Interestingly, M is a supporting matching since the mixed matching $\Pi = \{(M, \frac{1}{2}), (P, \frac{1}{2})\}$ is popular. Moreover, M is fairly popular since N is the only matching that defeats M and observe that N leaves the stable vertex a_1 unmatched, hence N is *not* a supporting matching.

A hardness result. As observed above, it is not the case that every unpopular matching has to be defeated by some popular matching. This motivates the following question: how easy is it to decide if there exists a popular matching that defeats a given matching M? This is a natural question when matching M is already in place and we want to replace M with a popular matching. An ideal matching would be a popular matching that is more popular than M, if such a matching exists. Interestingly, we can show a "compactness" result. Note that G may have more than 2^n popular matchings [32].

▶ **Proposition 7.** *There is a set K of at most m popular matchings in G such that any matching defeated by some popular matching in G has to be defeated by a matching in K.*

However, deciding if there is a popular matching that defeats a given matching is hard.

▶ **Theorem 8.** *Given a marriage instance $G = (A \cup B, E)$ and a matching M in G, it is NP-complete to decide if there exists any popular matching that is more popular than M.*

▶ Remark 9. It was mentioned earlier that it is coNP-hard to compute a min-cost matching that is not defeated by any popular matching. This hardness follows from Theorem 8 by setting $\mathsf{cost}(e) = 0$ for each $e \in M$ and $\mathsf{cost}(e) = 1$ for any $e \notin M$.

For any matching M, if there is a popular matching that defeats M then it is natural to regard M as a very unpopular matching (as there is a popular matching better than M). However to define a *mildly popular* matching as one that is undefeated by popular matchings would not have been very helpful as we know it is coNP-hard to identify such matchings (by Theorem 8). A natural strengthening of this property would have been to say that a matching M is mildly popular if and only if M is undefeated by popular *mixed* matchings. This is precisely one of the characterizations of supporting matchings (by Theorem 5).

Related results. The min-cost stable matching problem is very well-studied with several polynomial time algorithms [11, 12, 13, 20, 33] to solve this problem; furthermore, the stable matching polytope has a simple and elegant linear size formulation in $\mathbb{R}^m$ [29, 31]. It is known that the popular fractional matching polytope of G is half-integral [19].

A min-cost popular matching in G can be computed in $O^*(2^{n/4})$ time [25]. The intractability of the min-cost popular matching problem has motivated relaxations such as *quasi-popularity* [9] and *semi-popularity* [25]. A matching M is quasi-popular if $u(M) \leq 2$. Computing a min-cost quasi-popular matching is NP-hard; however a quasi-popular matching of cost at most that of a min-cost popular matching can be computed in polynomial time [9]. A matching M is semi-popular if $\Delta(M, N) \geq 0$ for at least half the matchings N in G. A bicriteria approximation algorithm was given in [25] to find an *almost* semi-popular matching whose cost is at most twice the cost of a min-cost popular matching.

Our techniques. The characterization of supporting matchings (given in Section 2) uses the half-integrality of the popular fractional matching polytope in a marriage instance [19] along with Hall's theorem. A technical lemma used here (and proved in the appendix) is based on the existence of certain helpful stable matchings as shown in [17].

Our characterization of supporting matchings implies that a matching M is fairly popular if and only if $M = \cup_C M_c$, where C is a connected component in the subgraph $(A \cup B, E_p)$ and every matching M_c in this decomposition has a certain *witness* or dual certificate. We show a surjective mapping from the union of sets of stable matchings in two auxiliary graphs G'_c and G''_c to the set of such matchings M_c. Let $\mathcal{S}'_c$ (resp., $\mathcal{S}''_c$) be the stable matching polytope of G'_c (resp., G''_c). The convex hull of $\mathcal{S}'_c \cup \mathcal{S}''_c$ is an extension of the convex hull of edge incidence vectors of such matchings M_c. Using Balas' theorem [2] to formulate the convex hull of $\mathcal{S}'_c \cup \mathcal{S}''_c$ leads to Theorem 4 proved in Section 3.

The LP-machinery for popular matchings was introduced in [26] and used in [19, 22] to study popular fractional matchings. The graphs G'_c and G''_c are inspired by instances from [7, 23, 24] that solve variants of the popular matching problem by modeling them as stable matching problems in appropriate graphs. Our novelty is in our characterization of supporting matchings – this leads to a characterization of fairly popular matchings which allows us to formulate an extension of the fairly popular matching polytope $\mathcal{F}$ with $\text{poly}(m, n)$ many constraints, i.e., we show the polytope $\mathcal{F}$ has a compact extended formulation.

Our NP-hardness proof (given in Section 4) is based on the NP-hardness (from [10]) of deciding if there exists a popular matching that contains a given pair of edges.

2 A Characterization of Supporting Matchings

We prove Theorem 5 in this section. Before we characterize supporting matchings, it will be useful to recall some properties of popular fractional matchings in a marriage instance G.

A fractional matching x in G is a convex combination of matchings (by Birkhoff-von Neumann theorem). Recall that x is popular if Π is a popular mixed matching, where Π is any mixed matching that is equivalent to x. Alternatively, as shown in [26], x is popular if $\Delta(x, M) \ge 0$ for all matchings M where $\Delta(x, M) = \sum_{u \in A \cup B} \mathsf{vote}_u(x, M)$ and $\mathsf{vote}_u(x, M)$ is u's fractional vote (a value in $[-1, 1]$) for its assignment in x versus its assignment in M. Section 4 has more details on comparing a matching M with a fractional matching x.

The popular fractional matching polytope of G is the convex hull of all popular fractional matchings in G. It was shown in [19] that the popular fractional matching polytope of G is half-integral. The proof of half-integrality uses the graph $H = (A_H \cup B_H, E_H)$ defined below.

The graph H can be regarded as consisting of *two* copies of $G = (A \cup B, E)$ (see Figure 1). The vertex set $A_H = A_0 \cup B_1$ and $B_H = B_0 \cup A_1$, where $A_i = \{a_i : a \in A\}$ and $B_i = \{b_i : b \in B\}$ for $i = 0, 1$. The edge set E_H of H is described below.

- For every $(a, b) \in E$, there are 2 edges (a_0, b_0) and (a_1, b_1) in E_H.
- For every $u \in A \cup B$, there is a single edge (u_0, u_1) in E_H.

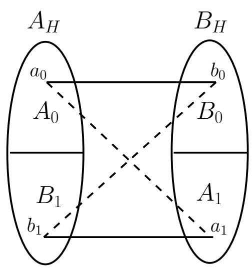

Figure 1 The vertex set of H has 2 copies u_0 and u_1 of every vertex u in G.

For any $u \in A \cup B$: if u's preference order in G is $v \succ v' \succ \cdots \succ v''$ then u_i's preference order (for $i = 0, 1$) in H is $v_i \succ v'_i \succ \cdots \succ v''_i \succ u_{1-i}$; so u_i's last choice neighbor is u_{1-i}.

The graph H admits a *perfect* stable matching, i.e., one that matches all vertices. Let S be any stable matching in G. Consider the matching S' in H defined as $S_0 \cup S_1 \cup \{(u_0, u_1) : u$ is unmatched in $S\}$ where $S_i = \{(a_i, b_i) : (a, b) \in S\}$ for $i = 0, 1$. It is easy to see that S' is a perfect stable matching in H.

It was shown in [19, Theorem 2] that if a marriage instance has a perfect stable matching then its popular fractional matching polytope is integral. Thus the popular fractional matching polytope of H is integral.

The function f. For any matching N in G, there is a corresponding matching N' in H defined as $\{(a_0, b_0), (a_1, b_1) : (a, b) \in N\} \cup \{(u_0, u_1) : u \text{ is unmatched in } N\}$. This map extends to fractional matchings, so for any fractional matching x in G, there is a corresponding fractional matching x' in H. Similarly, there is a map f from the set of fractional matchings in H to the set of fractional matchings in G: $f(y) = x$ where $x_{(a,b)} = (y_{(a_0,b_0)} + y_{(a_1,b_1)})/2$ for any $(a, b) \in E$. Observe that $f(x') = x$ where x' is the fractional matching in H that corresponds to x in G. If the fractional matching y is popular in H then the fractional matching $f(y)$ is popular in G since $\Delta(f(y), N) = \Delta(y, N')/2$ for any matching N in G.

Note that $(a, b) \in E$ is a popular fractional edge in G, i.e., $(a, b) \in E_p$, if and only if (a_0, b_0) and (a_1, b_1) are popular fractional edges in H. Since the popular fractional matching polytope of H is integral, it follows that (a_0, b_0) and (a_1, b_1) are *popular edges*[1] in H. Also, (u_0, u_1) is a popular edge in H if and only if u is an unstable vertex in G.

Proof of Theorem 5

We need to show the following three statements are equivalent.

1. M is supporting.
2. No popular mixed matching defeats M.
3. M matches all stable vertices and $M \subseteq E_p$.

Proof of 1⇒2. Let M be a supporting matching. Then there exists a popular mixed matching $\Pi = \{(M_0, p_0), \ldots, (M_k, p_k)\}$ where $M = M_i$ for some i. Suppose there is a popular mixed matching Π' that defeats M, i.e., $\Delta(\Pi', M) > 0$. Because both Π and Π' are popular mixed matchings, we have $\Delta(\Pi', \Pi) = \sum_j p_j \cdot \Delta(\Pi', M_j) = 0$. Since $\Delta(\Pi', M_i) > 0$ and $\Delta(\Pi', \Pi) = 0$, there has to exist some matching M_j on which Π has support such that $\Delta(\Pi', M_j) < 0$. However this contradicts Π''s popularity, thus 1⇒2.

Proof of 2⇒3. This part needs the following technical lemma. The proof of Lemma 10 uses the existence of certain helpful stable matchings as shown in [17] and is given in the appendix. Call an edge e *unpopular* if there exists no popular matching that contains e.

▶ **Lemma 10.** *Any matching in H that contains an unpopular edge is defeated by some popular matching in H.*

Let M be a matching in G such that either M has an edge not in E_p or some stable vertex is left unmatched in M. So the matching $M' = \{(a_0, b_0), (a_1, b_1) : (a, b) \in M\} \cup \{(u_0, u_1) : u$ is unmatched in $M\}$ in H has an edge that is not a popular edge. Then some popular matching P in H defeats M' (by Lemma 10).

Recall the map f from the set of fractional matchings in H to the set of fractional matchings in G defined earlier in Section 2. Let $r = f(P)$. The fractional matching r is popular in G because P is a popular matching in H. Since $\Delta(P, M') > 0$, we have $\Delta(r, M) > 0$. The fractional matching r can be regarded as a mixed matching Π; moreover, Π is popular since r is popular. Thus there is a popular mixed matching Π that is more popular than M, a contradiction to M satisfying property 2. Thus 2⇒3.

[1] An edge e is popular if there is a popular matching that contains e.

Proof of 3⇒1. Let $e = (a, b) \in M$. Since $M \subseteq E_p$, by what was discussed earlier in Section 2, there are popular matchings M_e^0 and M_e^1 in H that contain (a_0, b_0) and (a_1, b_1), respectively. For any vertex u left unmatched in M, it has to be the case that u is an unstable vertex in G. So there is a popular matching M_u in H that contains (u_0, u_1).

Suppose $M = \{e_1, \ldots, e_\ell\}$ and let $u_1, \ldots, u_t$ be left unmatched in M. Consider the $2\ell + t$ matchings $M_{e_1}^0, \ldots, M_{e_\ell}^0, M_{e_1}^1, \ldots, M_{e_\ell}^1$ and $M_{u_1}, \ldots, M_{u_t}$ in H analogous to the matchings M_e^0, M_e^1, and M_u defined above. Let H' be the graph whose edge set is the multiset of edges present in these $2\ell + t$ matchings, i.e., multiple copies of an edge are present in this edge set if this edge is present in more than one matching. The graph H' is $(2\ell + t)$-regular since each of these $2\ell + t$ matchings is popular and hence, perfect in H (recall that H has a perfect stable matching and stable matchings are min-size popular matchings).

Observe that $M' = \{(a_0, b_0), (a_1, b_1) : (a, b) \in M\} \cup \{(u_0, u_1) : u \text{ is unmatched in } M\}$ belongs to H'. Delete M' from H'. Since M' is a perfect matching in H', the resulting graph $H'' = H' \setminus M'$ is $(2\ell + t - 1)$-regular. It follows from Hall's theorem that H'' can be decomposed into $2\ell + t - 1$ perfect matchings $N'_1, \ldots, N'_{2\ell+t-1}$. Thus we have:

$$I_{M'} + I_{N'_1} + \cdots + I_{N'_{2\ell+t-1}} \;=\; I_{M_{e_1}^0} + \cdots + I_{M_{e_\ell}^1} + I_{M_{u_1}} + \cdots + I_{M_{u_t}},$$

where for any matching N, the vector I_N is its edge incidence vector.

The $2\ell + t$ matchings $M_{e_1}^0, \ldots, M_{e_\ell}^1, M_{u_1}, \ldots, M_{u_t}$ (on the right hand side above) are popular in H. Hence the fractional matching $q = (I_{M_{e_1}^0} + \cdots + I_{M_{u_t}})/(2\ell + t)$, which can also be written as $(I_{M'} + I_{N'_1} + \cdots + I_{N'_{2\ell+t-1}})/(2\ell + t)$, is popular in H.

So $r = f(q)$ is a popular fractional matching in G. The mixed matching $\Pi = \{(M, \frac{1}{2\ell+t}), \ldots\}$ is equivalent to r and it has support on M. Moreover, Π is a popular mixed matching since r is a popular fractional matching. Thus M is a supporting matching. Hence 3⇒1. ◀

3 The Fairly Popular Matching Polytope

We prove Theorem 4 in this section. We will see an LP framework for fairly popular matchings in Section 3.1. A characterization of fairly popular matchings will be given in Section 3.2. In Sections 3.3 and 3.4, this characterization will be used to solve the min-cost fairly popular matching problem in polynomial time.

3.1 An LP Framework

Our input instance is $G = (A \cup B, E)$. Let $E_p \subseteq E$ be the set of popular fractional edges in G. The set E_p can be computed in linear time by running the popular edge algorithm (from [7]) in the instance H described in Section 2.

Let $\tilde{E}_p = E_p \cup \{(u, u) : u \text{ is an unstable vertex in } G\}$ and let $G_p = (A \cup B, \tilde{E}_p)$. We know from Theorem 5 that every perfect matching $\tilde{N}$ in G_p is a supporting matching N augmented with self-loops at vertices left unmatched in N; conversely, every supporting matching N augmented with self-loops at unmatched vertices is a perfect matching $\tilde{N}$ in G_p.

Let M be any matching in G. In order to decide if there exists a supporting matching that defeats M, the following edge weight function in G_p will be useful. For any $(a, b) \in E_p$:

$$\text{let } \mathsf{wt}_M(a,b) = \begin{cases} 2 & \text{if } (a,b) \text{ is a blocking edge to } M; \\ -2 & \text{if } a \text{ and } b \text{ prefer their partners in } M \text{ to each other;} \\ 0 & \text{otherwise.} \end{cases}$$

For any unstable vertex u, let $\mathsf{wt}_M(u, u) = 0$ if u is left unmatched in M, else $\mathsf{wt}_M(u, u) = -1$.

Consider the following linear program (LP1). For any vertex v, let $\delta_p(v)$ be the set of edges incident to v in G_p.

$$\text{maximize} \sum_{e \in \tilde{E}_p} \mathsf{wt}_M(e) \cdot x_e \tag{LP1}$$

subject to

$$\sum_{e \in \delta_p(v)} x_e = 1 \quad \forall\, v \in A \cup B \qquad \text{and} \qquad x_e \geq 0 \quad \forall\, e \in \tilde{E}_p.$$

Since the constraint matrix is totally unimodular, (LP1) is integral. This LP computes a max-weight perfect matching $\tilde{N}$ in G_p (so N is supporting by Theorem 5) with respect to the edge weight function wt_M. The following claim is easy to see.

▷ Claim 11. For any perfect matching $\tilde{N}$ in G_p, we have $\mathsf{wt}_M(\tilde{N}) = \Delta(N, M)$.

Proof. For any edge $e = (a, b) \in E_p$, observe that $\mathsf{wt}_M(e) = \mathsf{vote}_a(b, M) + \mathsf{vote}_b(a, M)$ where for any vertex v and neighbor v', $\mathsf{vote}_v(v', M) \in \{\pm 1, 0\}$ is v's vote for v' versus its assignment in M. So $\mathsf{vote}_v(v', M) = 1$ if v prefers v' to its assignment in M, it is -1 if v prefers its assignment in M to v', otherwise it is 0. Similarly, for any unstable vertex v, $\mathsf{wt}_M(v, v)$ is 0 if M leaves v unmatched, else it is -1.

Hence for any perfect matching $\tilde{N}$ in G_p, observe that $\mathsf{wt}_M(\tilde{N})$ is the sum of votes of all vertices, where each vertex votes for its assignment in N versus its assignment in M. In other words, $\mathsf{wt}_M(\tilde{N}) = \phi(N, M) - \phi(M, N) = \Delta(N, M)$. ◁

It follows from Claim 11 that if the optimal value of (LP1) is positive then there exists a supporting matching that defeats M; else $\Delta(N, M) \leq 0$ for all supporting matchings N, so M is fairly popular. Note that for any stable matching N in G, we have $\mathsf{wt}_M(\tilde{N}) = \Delta(N, M) \geq 0$ (due to N's popularity in G). So the optimal value of (LP1) has to be at least 0. Hence M is fairly popular if and only if the optimal value of (LP1) is 0.

Let $U \subseteq A \cup B$ be the set of unstable vertices in G. The linear program (LP2) is the dual LP.

$$\text{minimize} \sum_{v \in A \cup B} \alpha_v \tag{LP2}$$

subject to

$$\alpha_a + \alpha_b \;\geq\; \mathsf{wt}_M(a, b) \;\; \forall\, (a, b) \in E_p \qquad \text{and} \qquad \alpha_u \;\geq\; \mathsf{wt}_M(u, u) \;\; \forall\, u \in U.$$

So M is fairly popular if and only if the optimal value of (LP2) is 0.

3.2 Witnesses for Fairly Popular Matchings

Let C be any connected component in $G_p = (A \cup B, \tilde{E}_p)$. Since all stable matchings in G match the stable vertices of C among themselves, the number of stable vertices in $C_A = C \cap A$ is the same as the number of stable vertices in $C_B = C \cap B$. Hence there are k stable vertices in C_A if and only if there are k stable vertices in C_B.

▶ **Lemma 12.** *A matching M is fairly popular if and only if there exists a feasible solution α to* (LP2) *such that for every connected component C in G_p, we have $\sum_{v \in C} \alpha_v = 0$ and furthermore,*

- *either $\alpha_v \in \{0, \pm 2, \pm 4, \ldots, \pm 2k\}$ for all $v \in C$*
- *or $\alpha_v \in \{\pm 1, \pm 3, \pm 5, \ldots, \pm(2k+1)\}$ for all $v \in C$,*

where $2k$ is the number of stable vertices in C.

Proof. Let M be a matching such that there exists a feasible solution α to (LP2) with $\sum_{v \in C} \alpha_v = 0$ for every connected component C in G_p. Then $\sum_{v \in A \cup B} \alpha_v = 0$ and so the optimal value of (LP2) is 0. Hence M is fairly popular.

Conversely, let M be a fairly popular matching in G and let α be an optimal solution to (LP2). The constraint matrix of (LP2) is totally unimodular, so we can assume that $\alpha \in \mathbb{Z}^n$.

Let C be any connected component in G_p. We have $\mathsf{wt}_M(\tilde{N}_c) \geq 0$ where N is any stable matching in G and $N_c = N \cap (C \times C)$. Hence $\sum_{v \in C} \alpha_v \geq 0$. Moreover, $\sum_C \sum_{v \in C} \alpha_v = \sum_{v \in A \cup B} \alpha_v = 0$ since M is fairly popular. Hence it has to be the case that $\sum_{v \in C} \alpha_v = 0$ for every connected component C in G_p.

Every edge in E_p belongs to some popular fractional matching in G. Let q be the popular fractional matching that $(a, b) \in E_p$ belongs to, where a and b are vertices in C. We have $\Delta(q, M) = 0$ since q is a popular fractional matching, thus q is an optimal solution to (LP1). Because α is an optimal solution to (LP2), we have $\alpha_a + \alpha_b = \mathsf{wt}_M(a, b)$ by complementary slackness, i.e., every edge in G_p is tight. So $\alpha_a + \alpha_b = \mathsf{wt}_M(a, b) \in \{0, \pm 2\}$ for all $(a, b) \in E_p$. Hence the α-values of all the vertices in C have the same parity.

Suppose every vertex of C is stable. Then we can update the α-values of vertices in C as follows for any value t: $\alpha_a = \alpha_a - t$ for all $a \in C_A$ and $\alpha_b = \alpha_b + t$ for all $b \in C_B$. The updated α-values are also a feasible solution to (LP2) since the sum $\alpha_a + \alpha_b$ for any $(a, b) \in E_p$ (where a and b are in C) is unchanged by this update; moreover, we assumed that C has no unstable vertex, so there is no constraint $\alpha_u \geq \mathsf{wt}_M(u, u)$ for any $u \in C$.

Moreover, the sum of α-values of all vertices in C is unchanged by this update since $|C_A| = |C_B| = k$ (because C has only stable vertices), so $\sum_{v \in C} \alpha_v = 0$. Thus we can preserve optimality and shift α-values so as to make $\alpha_v = 0$ for some $v \in C$. All the edges in G_p are tight, so the matched partners of vertices with α-value 0 also have α-value 0 and all neighbors in C of vertices with α-value 0 have their α-values in $\{0, \pm 2\}$. Their partners have α-values in $\{0, \pm 2\}$ and neighbors of these vertices have α-values in $\{0, \pm 2, \pm 4\}$ and so on. Since the number of stable vertices in C_A (and also in C_B) is k, we can conclude that there exists an optimal solution α to (LP2) such that $\alpha_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$.

Let us now assume that C has at least one unstable vertex. Consider the matching $\tilde{N} = N \cup \{(u, u) : u \in U\}$, where N is any stable matching in G and U is the set of unstable vertices in G. The matching $\tilde{N}$ is an optimal solution to (LP1). By complementary slackness, we have $\alpha_u = \mathsf{wt}_M(u, u)$ for every $u \in U$. Hence $\alpha_u \in \{0, -1\}$ for every $u \in U$. Since the α-values of all the vertices in C have the same parity, we have the following two cases.

Case 1. The α-values of all the vertices in C are even. Then $\alpha_u = 0$ for every $u \in U \cap C$. As argued above (when C had no unstable vertex), this implies that $\alpha_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$.

Case 2: The α-values of all the vertices in C are odd. Then $\alpha_u = -1$ for every $u \in U \cap C$. An analogous argument to the one above shows that $\alpha_v \in \{\pm 1, \pm 3, \ldots, \pm(2k + 1)\}$ for all $v \in C$. ◀

A characterization of fairly popular matchings. By Lemma 12, a matching M is fairly popular if and only if $M = \cup_C M_c$ where for every connected component C in G_p, there exists γ (this is the vector α in Lemma 12 restricted to vertices in C) such that:

1. $\sum_{v \in C} \gamma_v = 0$;
2. $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a, b)$ for $(a, b) \in E_p \cap (C \times C)$ and $\gamma_u \geq \mathsf{wt}_{M_c}(u, u)$ for $u \in U \cap C$;
3. either $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$ or $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k + 1)\}$ for all $v \in C$, where $2k$ is the number of stable vertices in C.

Witnesses. We know that M is fairly popular if and only if for each connected component C in G_p, there exists γ such that $M_c = M \cap (C \times C)$ and γ satisfy properties 1-3 given above. Such a vector γ will be called a *witness* of M_c. Let $G_c = (C, E_c)$ where $E_c = E_p \cap (C \times C)$.

▶ **Definition 13.** *Call a matching M_c in G_c* valid *if it has a witness, i.e., there exists a vector γ such that M_c and γ satisfy properties 1-3 given above.*

Let $\mathcal{F}_c$ be the convex hull of edge incidence vectors of all valid matchings in G_c. By Lemma 12, $\mathcal{F}_c$ is the convex hull of $\mathcal{F}_c^0 \cup \mathcal{F}_c^1$ where:

- $\mathcal{F}_c^0$ is the convex hull of edge incidence vectors of valid matchings in G_c with a witness γ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$.
- $\mathcal{F}_c^1$ is the convex hull of edge incidence vectors of valid matchings in G_c with a witness γ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k+1)\}$ for all $v \in C$.

3.3 Two Useful Stable Matching Instances

Let C be any connected component in G_p with $|C| \geq 2$. We will now describe instances G_c' and G_c'' such that the stable matching polytope of G_c' (resp., G_c'') is an extension of $\mathcal{F}_c^0$ (resp., $\mathcal{F}_c^1$). Let S be the set of stable vertices in G and let $|S \cap C| = 2k$.

The instance $G_c' = (A_c' \cup B_c', E_c')$. Every $a \in S \cap C_A$ has $2k+1$ copies $a_{-k}, \ldots, a_0, \ldots, a_k$ in A_c'. Recall that U is the set of unstable vertices in G. Every $a \in U \cap C_A$ has exactly one copy a_0 in A_c'.

Let $B_c' = \{\tilde{b} : b \in C_B\} \cup \{d_{1-k}(a), \ldots, d_k(a) : a \in S \cap C_A\}$, where the set $\{\tilde{b} : b \in C_B\}$ is a copy of C_B. Along with vertices in $\{\tilde{b} : b \in C_B\}$, the set B_c' contains $2k$ *dummy* vertices $d_{1-k}(a), \ldots, d_k(a)$ for each $a \in S \cap C_A$. The purpose of the $2k$ dummy vertices $d_{1-k}(a), \ldots, d_k(a)$ is to ensure that only one of $a_{-k}, \ldots, a_{-1}, a_0, a_1, \ldots, a_k$ is matched to a *non-dummy* neighbor in any stable matching in G_c'.

For any $a \in S \cap C_A$, the set E_c' has the edges $(a_{i-1}, d_i(a))$ and $(a_i, d_i(a))$ for $1-k \leq i \leq k$. For every edge (a, b) in E_c, the following edges are in E_c'. Since vertices in U form an independent set, note that at least one of a, b has to be in S.

1. If only one of a, b is in S then there is only one edge $(a_0, \tilde{b})$ in E_c'.
2. If both a and b are in S then there are $2k+1$ edges $(a_i, \tilde{b})$ in E_c' where $-k \leq i \leq k$.

Let a's preference order among its neighbors in G_c be $b_1 \succ \cdots \succ b_r$.

- If $a \in U$ then the preference order of a_0 is $\tilde{b}_1 \succ \cdots \succ \tilde{b}_r$.
- Suppose $a \in S$. The vertex a_0's preference order is $d_0(a) \succ \tilde{b}_1 \succ \cdots \succ \tilde{b}_r \succ d_1(a)$. Note that all of a's neighbors in G_c are present in a_0's preference list – this will not be so for a_i, where $i \neq 0$. Let $t_1, \ldots, t_s$ be a's neighbors in G_c that are in S. Let a's preference order among these neighbors be $t_1 \succ \cdots \succ t_s$.
  - a_{-k}'s preference order in G_c' is $\tilde{t}_1 \succ \cdots \succ \tilde{t}_s \succ d_{1-k}(a)$.
  - For $i \in \{1-k, \ldots, k-1\} \setminus \{0\}$: a_i's preference order is $d_i(a) \succ \tilde{t}_1 \succ \cdots \succ \tilde{t}_s \succ d_{i+1}(a)$.
  - a_k's preference order in G_c' is $d_k(a) \succ \tilde{t}_1 \succ \cdots \succ \tilde{t}_s$.

  For any i, the preference order of $d_i(a)$ is $a_{i-1} \succ a_i$.

Consider any $b \in C_B$. Let b's preference order for its neighbors in G_c be $a \succ \cdots \succ z$. If $b \in U$ then $\tilde{b}$'s preference order for its neighbors in G_c' is $a_0 \succ \cdots \succ z_0$.

Suppose $b \in S$. Let $\{a', \ldots, z'\} \subseteq \{a, \ldots, z\}$ be the set of b's neighbors in G_c that are in S. Let b's preference order among these neighbors be $a' \succ \cdots \succ z'$. The preference order of $\tilde{b}$ in G_c' is:

$$\underbrace{a_k' \succ \cdots \succ z_k'}_{\text{level } k \text{ neighbors}} \succ \cdots \succ \underbrace{a_1' \succ \cdots \succ z_1'}_{\text{level 1 neighbors}} \succ \underbrace{a_0 \succ \cdots \succ z_0}_{\text{level 0 neighbors}} \succ \cdots \succ \underbrace{a_{-k}' \succ \cdots \succ z_{-k}'}_{\text{level } -k \text{ neighbors}}$$

So copies of all neighbors of b in G_c are present only in level 0. Note that $\tilde{b}$ prefers subscript/level i neighbors to level j neighbors for any $i > j$.

Stable matchings in G'_c. For any valid matching M_c in G_c with a witness γ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$, define M'_c in G'_c as follows. For every $(a, b) \in M_c$:

- include the edge $(a_i, \tilde{b})$ in M'_c where $\gamma_a = -2i$;
- for $j < i$ and $a \in S$ do: add the edge $(a_j, d_{j+1}(a))$ to M'_c;
- for $j > i$ and $a \in S$ do: add the edge $(a_j, d_j(a))$ to M'_c.

We will show in Lemma 14 that M'_c is a stable matching in G'_c. Conversely, let M'_c be any stable matching in G'_c. Let M_c be the *preimage* of M'_c, i.e., M_c is obtained by deleting all edges in M'_c that are incident to dummy vertices and replacing any edge $(a_i, \tilde{b}) \in E'_c$ with $(a, b) \in E_c$. Note that M_c is a matching in G_c because all dummy vertices (being top choice neighbors) have to be matched in any stable matching in G'_c and so at most one of the a_i's can be matched to a non-dummy neighbor in M'_c.

We will show in Lemma 14 that M_c is a valid matching in G_c. The proof of Lemma 14 uses ideas from [23, 24] and is given in the appendix.

▶ **Lemma 14.** *M_c is a valid matching in G_c with a witness γ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$ if and only if M'_c is a stable matching in G'_c.*

The instance $G''_c = (A''_c \cup B''_c, E''_c)$. Every $a \in S \cap C_A$ has $2k + 2$ copies $a_{-k}, \ldots, a_{-1}$, $a_0, \ldots, a_{k+1}$ in A''_c. Every $a \in U \cap C_A$ has $k + 2$ copies $a_{-k}, \ldots, a_{-1}, a_0, a_1$ in A''_c. Let $B''_c = \{\tilde{b} : b \in C_B\} \cup \{d_{1-k}(a), \ldots, d_{k+1}(a) : a \in S \cap C_A\} \cup \{d_{1-k}(a), \ldots, d_1(a) : a \in U \cap C_A\}$. As before, the set $\{\tilde{b} : b \in C_B\}$ is a copy of the set C_B. Along with vertices in $\{\tilde{b} : b \in C_B\}$, the set B''_c contains $2k + 1$ dummy vertices for each $a \in S \cap C_A$ and $k + 1$ dummy vertices for each $a \in U \cap C_A$. For each edge $(a, b) \in E_c$, the following edges are in E''_c:

1. If $a \in U$ (so $b \in S$) then there are $k + 2$ edges $(a_i, \tilde{b})$ in E''_c where $-k \le i \le 1$.
2. If $b \in U$ (so $a \in S$) then there are $k + 2$ edges $(a_i, \tilde{b})$ in E''_c where $0 \le i \le k + 1$.
3. If both a and b are in S then there are $2k + 2$ edges $(a_i, \tilde{b})$ in E''_c where $-k \le i \le k + 1$.

Let $a \in C_A$. The set E''_c also has the edges $(a_{i-1}, d_i(a))$ and $(a_i, d_i(a))$ for $1-k \le i \le k+1$ if $a \in S$ and for $1 - k \le i \le 1$ if $a \in U$. For any i, the preference order of $d_i(a)$ is $a_{i-1} \succ a_i$.

Let a's preference order among its neighbors in G_c be $b_1 \succ \cdots \succ b_r$. Let $t_1, \ldots, t_s$ be a's neighbors in G_c that are in S and let $t_1 \succ \cdots \succ t_s$ be a's preference order among these neighbors.

- a_{-k}'s preference order in G''_c is $\tilde{t}_1 \succ \cdots \succ \tilde{t}_s \succ d_{1-k}(a)$.
- a_i's preference order is $d_i(a) \succ \tilde{t}_1 \succ \cdots \succ \tilde{t}_s \succ d_{i+1}(a)$ for $1 - k \le i \le -1$.
- If $a \in U$ then all of a's neighbors are in S and a_0's preference order is $d_0(a) \succ \tilde{b}_1 \succ \cdots \succ \tilde{b}_r \succ d_1(a)$ and a_1's preference order is $d_1(a) \succ \tilde{b}_1 \succ \cdots \succ \tilde{b}_r$.
- If $a \in S$ then a_i's preference order is $d_i(a) \succ \tilde{b}_1 \succ \cdots \succ \tilde{b}_r \succ d_{i+1}(a)$ for $0 \le i \le k$ and a_{k+1}'s preference order is $d_{k+1}(a) \succ \tilde{b}_1 \succ \cdots \succ \tilde{b}_r$.

Consider any $b \in C_B$. Let b's preference order for its neighbors in G_c be $a \succ \cdots \succ z$. If $b \in U$ then $a, \ldots, z$ are in S and $\tilde{b}$'s preference order among its neighbors in G''_c is:

$$\underbrace{a_{k+1} \succ \cdots \succ z_{k+1}}_{\text{level } k+1 \text{ neighbors}} \succ \underbrace{a_k \succ \cdots \succ z_k}_{\text{level } k \text{ neighbors}} \succ \cdots \succ \underbrace{a_1 \succ \cdots \succ z_1}_{\text{level 1 neighbors}} \succ \underbrace{a_0 \succ \cdots \succ z_0}_{\text{level 0 neighbors}}$$

Suppose $b \in S$. Let $a', \ldots, z'$ be b's neighbors in G_c that are in S and let b's preference order among these neighbors be $a' \succ \cdots \succ z'$. Then the preference order of $\tilde{b}$ in G''_c is:

$$\underbrace{a'_{k+1} \succ \cdots \succ z'_{k+1}}_{\text{level } k+1 \text{ neighbors}} \succ \cdots \succ \underbrace{a'_2 \succ \cdots \succ z'_2}_{\text{level 2 neighbors}} \succ \underbrace{a_1 \succ \cdots \succ z_1}_{\text{level 1 neighbors}} \succ \cdots \succ \underbrace{a_{-k} \succ \cdots \succ z_{-k}}_{\text{level } -k \text{ neighbors}}$$

Note that copies of all neighbors of b in G_c are present in level i only for $-k \le i \le 1$.

Stable matchings in G''_c. For any valid matching M_c in G_c with a witness γ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k+1)\}$ for all $v \in C$, define M''_c in G''_c as follows. For every $(a, b) \in M_c$:

- include the edge $(a_i, \tilde{b})$ in M''_c where $\gamma_a = -(2i-1)$;
- for $j < i$ do: add the edge $(a_j, d_{j+1}(a))$ to M''_c;
- for $j > i$ do: add the edge $(a_j, d_j(a))$ to M''_c.

We will show that M''_c is a stable matching in G''_c. Conversely, let M''_c be any stable matching in G''_c. As before, let M_c be the preimage of M''_c; observe that M_c is a matching in G_c. Lemma 15 (proved in the appendix) shows that M_c is a valid matching in G_c.

▶ **Lemma 15.** *M_c is a valid matching in G_c with a witness γ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k+1)\}$ for all $v \in C$ if and only if M''_c is a stable matching in G''_c.*

3.4 A compact extended formulation

For any vertex v in G'_c, let $\delta'_c(v)$ be the set of edges incident to v in G'_c and for any neighbor u of v, let $\{w \succ_v u\}$ be the set of all neighbors of v in G'_c that v prefers to u. Let T'_c be the set of vertices in G'_c matched in any stable matching in this graph. Consider constraints (1)-(3) in variables y_e where $e \in E'_c$ and λ_c (this variable will be defined later).

$$\sum_{w:\, w \succ_{a_i} \tilde{b}} y_{(a_i, w)} + \sum_{s:\, s \succ_{\tilde{b}} a_i} y_{(s, \tilde{b})} + y_{(a_i, \tilde{b})} \quad \ge \quad \lambda_c \quad \forall (a_i, \tilde{b}) \in E'_c \tag{1}$$

$$\sum_{e \in \delta'_c(v)} y_e \quad \le \quad \lambda_c \quad \forall v \in A'_c \cup B'_c \tag{2}$$

$$\sum_{e \in \delta'_c(v)} y_e \; = \; \lambda_c \quad \forall v \in T'_c \qquad \text{and} \qquad y_e \quad \ge \quad 0 \quad \forall e \in E'_c. \tag{3}$$

Constraints (1)-(3) with 1 replacing λ_c (wherever λ_c occurs) describe the stable matching polytope $\mathcal{S}'_c$ of G'_c (by [29]). The stability constraint for any edge $(a_i, \tilde{b})$ in E'_c is given by (1) with 1 replacing λ_c. The stability constraint for edge $(a_{i-1}, d_i(a))$ (resp., $(a_i, d_i(a))$) is given by $\sum_{e \in \delta'_c(v)} y_e = 1$ with $v = a_{i-1}$ (resp., $v = d_i(a)$). Note that both a_{i-1} and $d_i(a)$ are in T'_c.

By Lemma 14, the constraints formulating $\mathcal{S}'_c$ along with $y_{(a,b)} = \sum_i y_{(a_i, \tilde{b})}$ for $(a, b) \in E_c$ describe an extension of the convex hull $\mathcal{F}^0_c$ of the edge incidence vectors of valid matchings in G_c with a witness γ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$.

For any vertex v in G''_c, let $\delta''_c(v)$ be the set of edges incident to v in G''_c and for any neighbor u of v, let $\{w \succ_v u\}$ be the set of all neighbors of v in G''_c that v prefers to u. Let T''_c be the set of vertices in G''_c matched in any stable matching in this graph. Consider constraints (4)-(6) in variables z_e where $e \in E''_c$ and λ_c.

$$\sum_{w:\, w \succ_{a_i} \tilde{b}} z_{(a_i, w)} + \sum_{s:\, s \succ_{\tilde{b}} a_i} z_{(s, \tilde{b})} + z_{(a_i, \tilde{b})} \quad \ge \quad 1 - \lambda_c \quad \forall (a_i, \tilde{b}) \in E''_c \tag{4}$$

$$\sum_{e \in \delta''_c(v)} z_e \quad \le \quad 1 - \lambda_c \quad \forall v \in A''_c \cup B''_c \tag{5}$$

$$\sum_{e \in \delta''_c(v)} z_e \; = \; 1 - \lambda_c \quad \forall v \in T''_c \qquad \text{and} \qquad z_e \quad \ge \quad 0 \qquad \forall e \in E''_c \tag{6}$$

Constraints (4)–(6) with 1 replacing $1-\lambda_c$ (wherever $1-\lambda_c$ occurs) describe the stable matching polytope $\mathcal{S}''_c$ of G''_c (by [29]). The stability constraint for $(a_i, \tilde{b}) \in E''_c$ is given by (4) with 1 replacing $1-\lambda_c$; the stability constraint for edge $(a_{i-1}, d_i(a))$ (resp., $(a_i, d_i(a))$) is given by $\sum_{e \in \delta''_c(v)} z_e = 1$ with $v = a_{i-1}$ (resp., $v = d_i(a)$). Both a_{i-1} and $d_i(a)$ are in T''_c.

By Lemma 15, the constraints formulating $\mathcal{S}''_c$ along with $z_{(a,b)} = \sum_i z_{(a_i,\tilde{b})}$ for $(a,b) \in E_c$ describe an extension of the convex hull $\mathcal{F}^1_c$ of the edge incidence vectors of valid matchings in G_c with a witness γ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k+1)\}$ for all $v \in C$.

We know from Lemma 12 that any valid matching in C has a witness γ where either (i) $\gamma_v \in \{0, \ldots, \pm 2k\}$ for all $v \in C$ or (ii) $\gamma_v \in \{\pm 1, \ldots, \pm(2k+1)\}$ for all $v \in C$. So the convex hull of $\mathcal{F}^0_c \cup \mathcal{F}^1_c$ is the valid matching polytope $\mathcal{F}_c$ of G_c. Consider constraints (7)-(8).

$$x_{(a,b)} = \sum_i y_{(a_i,\tilde{b})} + \sum_i z_{(a_i,\tilde{b})} \quad \forall (a,b) \in E_c \tag{7}$$

$$x_e = 0 \ \ \forall e \in (E \cap (C \times C)) \setminus E_c \qquad \text{and} \qquad 0 \le \lambda_c \le 1 \tag{8}$$

The summations over i in constraint (7) are over appropriate i, i.e., if a and b are in S then $x_{(a,b)} = \sum_{i=-k}^{k} y_{(a_i,\tilde{b})} + \sum_{i=-k}^{k+1} z_{(a_i,\tilde{b})}$. If a is in U then $x_{(a,b)} = y_{(a_0,\tilde{b})} + \sum_{i=-k}^{1} z_{(a_i,\tilde{b})}$ and if b is in U then $x_{(a,b)} = y_{(a_0,\tilde{b})} + \sum_{i=0}^{k+1} z_{(a_i,\tilde{b})}$.

Using Balas' theorem [2] to formulate an extension of the convex hull of $\mathcal{F}^0_c \cup \mathcal{F}^1_c$ introduces the variable $\lambda_c \in [0,1]$ and we get constraints (1)-(8) as given above. Thus the polytope defined by (1)-(8) is an extension of the polytope $\mathcal{F}_c$. Hence Theorem 16 follows.

▶ **Theorem 16.** *The polytope $\mathcal{P}_c$ defined by constraints* (1)-(8) *is an extension of the convex hull $\mathcal{F}_c$ of edge incidence vectors of valid matchings in G_c.*

For any two distinct connected components C and C' in G_p, the variables in the formulation of $\mathcal{P}_c$ and those in the formulation of $\mathcal{P}_{c'}$ are distinct. By listing the constraints in the formulation of $\mathcal{P}_c$ over all the non-trivial connected components C in G_p (i.e., $|C| \ge 2$) along with $x_e = 0$ for $e \in E \setminus \cup_C E_c$ (where the union is over all the non-trivial connected components C in G_p), we obtain a compact extended formulation for the fairly popular matching polytope of G. Linear programming on this formulation finds a min-cost fairly popular matching in G in polynomial time. This proves Theorem 4 stated in Section 1.

4 A Hardness Result

We prove Proposition 7 and Theorem 8 in this section. Let $\mathcal{M}_G$ be the matching polytope of the bipartite graph $G = (A \cup B, E)$ where $|A \cup B| = n$ and $|E| = m$. The polytope $\mathcal{M}_G \subseteq \mathbb{R}^m$ is described by the following constraints:

$$\sum_{e \in \delta(v)} x_e \le 1 \ \ \forall v \in A \cup B \qquad \text{and} \qquad x_e \ge 0 \ \ \forall e \in E.$$

For any vertex v, $\delta(v)$ is the set of edges in E incident to v. Any point $x \in \mathcal{M}_G$ is a fractional matching. Let $\tilde{E} = E \cup \{(v,v) : v \in A \cup B\}$ and let $\tilde{G} = (A \cup B, \tilde{E})$. That is, $\tilde{G}$ has self-loops (v,v) for all $v \in A \cup B$. The interpretation is that every vertex v is its own last choice neighbor. So we can regard any fractional matching x as a *perfect* fractional matching in $\tilde{G}$ by setting $x_{(v,v)} = 1 - \sum_{e \in \delta(v)} x_e$ for all vertices v.

For any matching M, recall the edge weight function wt_M defined in Section 3. This was defined in the graph $G_p = (A \cup B, \tilde{E}_p)$ and it easily extends (by the same definition) to $\tilde{G} = (A \cup B, \tilde{E})$. For any edge $e \in E$, $\mathsf{wt}_M(e) \in \{0, \pm 2\}$ and for any self-loop (v,v), $\mathsf{wt}_M(v,v) \in \{0, -1\}$. For any fractional matching x:

$$\Delta(x, M) = \mathsf{wt}_M(x) = \sum_{e \in \tilde{E}} \mathsf{wt}_M(e) \cdot x_e.$$

As shown in [26], this is exactly the same as defining $\Delta(x, M) = \Delta(\Pi, M)$ where Π is any mixed matching that is equivalent to x. Any popular matching M satisfies $\Delta(x, M) \le 0$ for all $x \in \mathcal{M}_G$. Note that the constraint $\Delta(x, M) \le 0$ involves $m + n$ variables x_e for $e \in \tilde{E}$. By substituting $x_{(v,v)} = 1 - \sum_{e\in\delta(v)} x_e$ for every vertex v, this constraint involves only the m variables x_e for $e \in E$.

▶ **Observation 17.** *Let $\mathcal{X} \subseteq \mathbb{R}^m$ be the convex hull of the edge incidence vectors of matchings that are not defeated by any popular matching. The polytope $\mathcal{X}$ is a face of $\mathcal{M}_G$.*

Proof. Every $x \in \mathcal{M}_G$ satisfies $\Delta(x, N) \le 0$ for all popular matchings N. So the intersection of $\mathcal{M}_G$ with the constraints $\Delta(x, N) = 0$ for all popular matchings N is a face $\mathcal{Q}$ of $\mathcal{M}_G$. The polytope $\mathcal{Q}$ is integral and every integral point in $\mathcal{Q}$ is the edge incidence vector of a matching not defeated by any popular matching. Moreover, the edge incidence vector of every matching that is not defeated by any popular matching is in $\mathcal{Q}$. Hence $\mathcal{Q} = \mathcal{X}$. ◀

The following constraints in the variables x_e for $e \in E$ describe the polytope $\mathcal{X}$:

$$\Delta(x, N) \;=\; 0 \;\forall \text{ popular matchings } N, \quad \sum_{e\in\delta(v)} x_e \le 1 \quad \forall\, v \in A \cup B, \quad \text{and} \quad x_e \;\ge\; 0 \;\forall\, e \in E.$$

There are exponentially many constraints here. However, $\mathcal{X}$ is a polytope in $\mathbb{R}^m$ and so at most m of the tight constraints $\Delta(x, N) = 0$ are necessary and the rest are redundant. Thus there exist at most $k \le m$ popular matchings $N_1, \ldots, N_k$ such that if a matching M satisfies $\Delta(M, N_i) = 0$ for $1 \le i \le k$ then the edge incidence vector of M belongs to $\mathcal{X}$, i.e., such a matching M is not defeated by any popular matching. Hence Proposition 7 follows.

The NP-hardness proof. We now prove Theorem 8 which states that in spite of the compactness result given by Proposition 7, it is NP-complete to decide if there exists a popular matching that defeats a given matching M. The reduction is from 1-in-3 SAT. This is the set of 3CNF formulas where each clause has 3 literals, none negated, such that there is a satisfying assignment that makes exactly one literal true in each clause.

Given such an input formula ψ, to decide if ψ is 1-in-3 satisfiable is NP-complete [30]. Given ψ, as done in [10], we will construct an instance G described below. The graph G has several gadgets. We are interested in two particular gadgets illustrated in Figure 2. These are on the 8 vertices: $a_0, z', u_0, u'_0 \in A$ and $b_0, z, v_0, v'_0 \in B$.

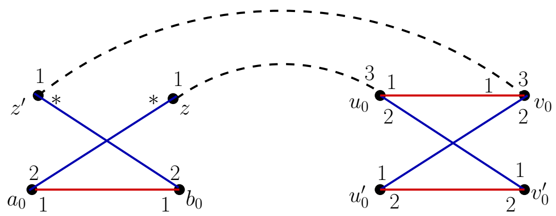

Figure 2 The numbers on edges denote preferences: 1 is top choice, 2 is second choice, and 3 is third choice; $*$ denotes a number > 1. The red edges are present in all stable matchings and the blue edges are present in all max-size popular matchings in G.

The top choices of z and z' are u_0 and v_0, respectively. However (z, u_0) and (z', v_0) (the dashed edges in Figure 2) do not belong to any popular matching. The vertices z, z' are adjacent to many vertices in the rest of the graph: we refer to [10] for these details – it is these vertices in the rest of the graph that represent the given formula ψ.

Let P be any popular matching in G. It was shown in [10] that P contains either (a_0, b_0) or the pair $(a_0, z), (z', b_0)$. Also P contains either the pair $(u_0, v_0), (u'_0, v'_0)$ or the pair $(u_0, v'_0), (u'_0, v_0)$. No other edge incident to any of these 8 vertices in Figure 2 belongs to any popular matching in G. The following hardness result [10, Theorem 4.2] will be crucial.

▶ **Theorem 18** ([10]). *The instance G has a popular matching that contains the three edges $(u_0, v'_0), (u'_0, v_0)$, and (a_0, b_0) if and only if ψ is 1-in-3 satisfiable.*

We will use the above instance $G = (A \cup B, E)$ to show the NP-hardness of deciding if there exists a popular matching that defeats a given matching M. Let $M = M_0 \cup M_1$ where $M_1 = \{(a_0, b_0), (u_0, z), (z', v_0), (u'_0, v'_0)\}$ and M_0 is any stable matching in the subgraph induced on $(A \cup B) \setminus S$, where $S = \{a_0, b_0, z', z, u_0, v_0, u'_0, v'_0\}$.

▶ **Lemma 19.** *There exists a popular matching in G that defeats M if and only if ψ is 1-in-3 satisfiable.*

Proof. Let G_1 be the subgraph of G induced on S and let G_0 be the subgraph induced on $(A \cup B) \setminus S$, where $S = \{a_0, b_0, z, z', u_0, v_0, u_1, v_1\}$.

The $\Rightarrow$ direction. Suppose there is a popular matching N that is more popular than M. No edge between G_0 and G_1 belongs to any popular matching [10], hence $N = N_0 \cup N_1$, where N_i is within G_i, for $i = 0, 1$. Since N is popular in G, the matchings N_0 and N_1 have to be popular in G_0 and G_1, respectively.

We have $\Delta(N, M) = \Delta(N_0, M_0) + \Delta(N_1, M_1)$. Since $\Delta(N, M) > 0$ and $\Delta(N_0, M_0) = 0$ (because M_0 and N_0 are popular matchings in G_0), it follows that $\Delta(N_1, M_1) > 0$.

The graph G_1 has three popular matchings and only one of them defeats M_1. This is the matching $\{(u_0, v'_0), (u'_0, v_0), (a_0, b_0)\}$ that leaves z, z' unmatched. It is easy to check that the other popular matchings in G_1 – these are $P = \{(a_0, b_0), (u_0, v_0), (u'_0, v'_0)\}$ and $P' = \{(a_0, z), (z', b_0), (u_0, v'_0), (u'_0, v_0)\}$ – do not defeat M_1.

So $N_1 = \{(u_0, v'_0), (u'_0, v_0), (a_0, b_0)\}$. We have $\Delta(N_1, M_1) = 4 - 2 = 2$ since u_0, v_0, u'_0, v'_0 prefer N_1 to M_1 while z, z' prefer M_1 to N_1 and a_0, b_0 are indifferent between N_1 and M_1. Since $N_1 \subseteq N$, it follows that N is a popular matching in G that contains $(u_0, v'_0), (u'_0, v_0)$, and (a_0, b_0). This means that ψ is 1-in-3 satisfiable (by Theorem 18).

The $\Leftarrow$ direction. Suppose ψ is 1-in-3 satisfiable. Then we know from Theorem 18 that there is a popular matching P that contains the edges $(u_0, v'_0), (u'_0, v_0), (a_0, b_0)$. We claim that $\Delta(P, M) > 0$. Let us partition P into $P_0 \cup P_1$ where $P_1 = \{(u_0, v'_0), (u'_0, v_0), (a_0, b_0)\}$ and $P_0 = P \setminus P_1$. We have $\Delta(P, M) = \Delta(P_1, M_1) + \Delta(P_0, M_0)$.

Observe that $\Delta(P_1, M_1) = 4 - 2 = 2$. Moreover, $\Delta(P_0, M_0) = 0$ by the popularity of P_0 and M_0 in G_0. So $\Delta(P, M) = 2$, i.e., the popular matching P defeats M. ◀

Lemma 19 shows that it is NP-hard to decide if there exists a popular matching that defeats a given matching M. This problem is NP-complete since a “yes”-instance M has a popular matching (which is easy to verify [4, 18]) that defeats it. Thus Theorem 8 stated in Section 1 follows.

5 Conclusions

We introduced a relaxation of popular matchings called *fairly popular* matchings in a marriage instance $G = (A \cup B, E)$. Unlike popular matchings, fairly popular matchings may lose to other matchings; however any matching N that defeats a fairly popular matching M does not belong to the support of any popular mixed matching, thus such a matching N can be considered to be quite *far* from being popular. So there is no "viable alternative" that defeats a fairly popular matching. Hence fairly popular matchings are a meaningful generalization of popular matchings.

We characterized matchings that belong to the support of popular mixed matchings. We showed that a matching M belongs to the support of a popular mixed matching if and only if M is undefeated by popular mixed matchings. We also gave a combinatorial characterization of such matchings. This allowed us to characterize fairly popular matchings in terms of witnesses and to use the stable matching machinery to formulate a compact extension of the fairly popular matching polytope. Thus the min-cost fairly popular matching problem can be solved in polynomial time. We also showed that it is NP-complete to decide if there exists a popular matching that is more popular than a given matching M.

References

1. A. Abdulkadiroğlu and T. Sönmez. School choice: a mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.
2. E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.
3. S. Baswana, P. P. Chakrabarti, S. Chandran, Y. Kanoria, and U. Patange. Centralized admissions for engineering colleges in India. *INFORMS Journal on Applied Analytics*, 49(5):338–354, 2019.
4. P. Biro, R. W. Irving, and D. F. Manlove. Popular matchings in the marriage and roommates problems. In *Proceedings of the 7th International Conference on Algorithms and Complexity (CIAC)*, pages 97–108, 2010.
5. Canadian Resident Matching Service. How the matching algorithm works. `http://carms.ca/algorithm.htm`.
6. Á. Cseh. Popular matchings. Trends in Computational Social Choice, Ulle Endriss (ed.), 2017.
7. Á. Cseh and T. Kavitha. Popular edges and dominant matchings. *Mathematical Programming*, 172(1):209–229, 2018.
8. L. Ehlers and T. Morrill. (Il)legal assignments in school choice. *The Review of Economic Studies*, 87(4):1837–1875, 2020.
9. Y. Faenza and T. Kavitha. Quasi-popular matchings, optimality, and extended formulations. In *Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 325–344, 2020.
10. Y. Faenza, T. Kavitha, V. Powers, and X. Zhang. Popular matchings and limits to tractability. In *Proceedings of the 30th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2790–2809, 2019.
11. T. Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45(2):233–284, 1992.
12. T. Feder. Network flow and 2-satisfiability. *Algorithmica*, 11(3):291–319, 1994.
13. T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28(1):103–126, 2003.
14. D. Gale and L.S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.
15. D. Gale and M. Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11(3):223–232, 1985.
16. P. Gärdenfors. Match making: assignments based on bilateral preferences. *Behavioural Science*, 20:166–173, 1975.
17. D. Gusfield and R. W. Irving. *The stable marriage problem: Structure and algorithms.* MIT Press, 1989.
18. C.-C. Huang and T. Kavitha. Popular matchings in the stable marriage problem. *Information and Computation*, 222:180–194, 2013.

19 C.-C. Huang and T. Kavitha. Popularity, mixed matchings, and self-duality. *Mathematics of Operations Research*, 46(2):405–427, 2021.

20 R. W. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the “optimal” stable marriage. *Journal of the ACM*, 34(3):532–543, 1987.

21 T. Kavitha. A size-popularity tradeoff in the stable marriage problem. *SIAM Journal on Computing*, 43(1):52–71, 2014.

22 T. Kavitha. Popular half-integral matchings. In *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 22:1–22:13, 2016.

23 T. Kavitha. Matchings, critical nodes, and popular solutions. In *Proceedings of the the 41st Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 25:1–25:19, 2021.

24 T. Kavitha. Maximum matchings and popularity. In *Proceedings of the 48th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 85:1–85:21, 2021.

25 T. Kavitha. Min-cost popular matchings. In *Proceedings of the the 40th Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 25:1–25:17, 2021.

26 T. Kavitha, J. Mestre, and M. Nasre. Popular mixed matchings. *Theoretical Computer Science*, 412:2679–2690, 2011.

27 M. McCutchen. The least-unpopularity-factor and least-unpopularity-margin criteria for matching problems with one-sided preferences. In *Proceedings of the 8th Latin American Symposium on Theoretical Informatics (LATIN)*, pages 593–604, 2008.

28 National Resident Matching Program. Why the Match? `http://www.nrmp.org/whythematch.pdf`.

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

30 T. J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing (STOC)*, pages 216–226, 1978.

31 C.-P. Teo and J. Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23(4):874–891, 1998.

32 E. G. Thurber. Concerning the maximum number of stable matchings in the stable marriage problem. *Discrete Mathematics*, 248(1-3):195–219, 2002.

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

A Appendix: Missing Proofs

We prove Lemma 10 here. Before we prove this lemma, we discuss some preliminaries that will be used in this proof. Let $\tilde{G} = (A \cup B, \tilde{E})$ be the graph G augmented with self-loops at all vertices. So each vertex v regards itself as its last choice neighbor and any matching M in G becomes a perfect matching $\tilde{M}$ in $\tilde{G}$ by augmenting M with self-loops at vertices left unmatched in M.

For any matching M, recall the edge weight function wt_M defined at the start of Section 3 in G_p. We now extend this edge weight function to all edges e in G, so $\mathsf{wt}_M(e) \in \{\pm 2, 0\}$ where $\mathsf{wt}_M(e) = 2$ if e blocks M and so on. For any vertex v, let $\mathsf{wt}_M(v,v) = 0$ if v is left unmatched in M, else $\mathsf{wt}_M(v,v) = -1$.

For any matching N in G, we have $\mathsf{wt}_M(\tilde{N}) = \Delta(N, M)$. So M is popular in G if and only if $\mathsf{wt}_M(\tilde{N}) \le 0$ for all matchings N. Since $\mathsf{wt}_M(\tilde{M}) = 0$, the matching M is popular in G if and only if the optimal value of (LP3) is 0. The linear program (LP4) is the dual LP.

▶ **Theorem 20** ([22, 26]). *A matching M in $G = (A \cup B, E)$ is popular if and only if there exists $y \in \{0, \pm 1\}^n$ such that $\sum_{v \in A \cup B} y_v = 0$ along with $y_a + y_b \ge \mathsf{wt}_M(a,b)$ for all $(a,b) \in E$ and $y_v \ge \mathsf{wt}_M(v,v)$ for all $v \in A \cup B$.*

$$\max \sum_{e \in \tilde{E}} \mathsf{wt}_M(e) \cdot x_e \qquad \text{(LP3)}$$
$$\text{s.t.} \quad \sum_{e \in \delta(v) \cup \{(v,v)\}} x_e = 1 \quad \forall\, v \in A \cup B$$
$$x_e \geq 0 \quad \forall\, e \in \tilde{E}.$$

$$\min \sum_{v \in A \cup B} y_v \qquad \text{(LP4)}$$
$$\text{s.t.} \quad y_a + y_b \geq \mathsf{wt}_M(a,b) \quad \forall\, (a,b) \in E$$
$$y_v \geq \mathsf{wt}_M(v,v) \quad \forall\, v \in A \cup B.$$

We will call a vector y, as given in Theorem 20, a *dual certificate* for popular matching M. Note that 0 is a dual certificate for any stable matching since a matching M is stable if and only if $\mathsf{wt}_M(e) \leq 0$ for all edges e.

We need to show in Lemma 10 that any matching in $H = (A_H \cup B_H, E_H)$ (see Figure 1) not defeated by any popular matching contains only popular edges. Every popular matching in H is perfect (since H has a perfect stable matching). As shown in [7], in such a case, there is a surjective map from the set of stable matchings in an auxiliary instance $H' = (A'_H \cup B'_H, E'_H)$ to the set of popular matchings in H. The graph H' is defined as follows:

- $A'_H = \{a, a' : a \in A_H\}$. So every $a \in A_H$ has two copies a and a' in A'_H.
- $B'_H = B_H \cup \{d(a) : a \in A_H\}$. So for every $a \in A_H$, there is a dummy vertex $d(a)$ in B'_H.

The vertex $d(a)$ has only two neighbors a, a' and $d(a)$ prefers a to a'. Every $(a,b) \in E_H$ has two copies (a,b) and (a',b) in E'_H. For any $a \in A_H$, if a's preference order in H is $b_1 \succ \cdots \succ b_r$ then a's preference order in H' is $b_1 \succ \cdots \succ b_r \succ d(a)$ and a''s preference order in H' is $d(a) \succ b_1 \succ \cdots \succ b_r$.

Let $b \in B_H$. If b's preference order in H is $a_1 \succ \cdots \succ a_k$ then b's preference order in H' is $a'_1 \succ \cdots \succ a'_k \succ a_1 \succ \cdots \succ a_k$, i.e., all its primed neighbors followed by all its unprimed neighbors, where the order among primed/unprimed neighbors is b's original order in H.

Let M' be any stable matching in H'. Then M' maps to the following matching in H: $M = \{(a,b) : (a,b) \text{ or } (a',b) \text{ is in } M'\}$.

For each $a \in A_H$, note that the stable matching M' has to match one of a, a' to $d(a)$ since $d(a)$ is the top choice neighbor for a'. The matching M is popular in H since it has the following witness $y \in \{\pm 1\}^{n_H}$: (where $|A_H \cup B_H| = n_H$)

1. for $a \in A_H$: if $(a', d(a)) \in M'$ then $y_a = 1$; else $y_a = -1$.
2. for $b \in B_H$: if b's partner in M' is a *primed* vertex (such as a') then $y_b = 1$; else $y_b = -1$.

We refer to [7, 19] for the details that y is a feasible solution to (LP2) and $\sum_{v \in A_H \cup B_H} y_v = 0$.

Proof of Lemma 10. Let N be a matching in H that contains an *unpopular* edge (s,t). We will now show there is a popular matching in H that defeats N. Call an edge e *stable* if there is a stable matching in H that contains e. The following result on stable matchings in a marriage instance will be useful to us.

▶ **Proposition 21** ([17, proof of Lemma 2.5.1])**.** *Suppose (s,t_0) and (s,t_1) are stable edges while (s,t) is not a stable edge where $t_1 \succ_s t \succ_s t_0$. Then there is a stable matching M where both s and t prefer their respective partners in M to each other.*

Let t_ℓ be the partner of s in the A_H-optimal stable matching M_ℓ in H and let t_r be the partner of s in the B_H-optimal stable matching M_r in H.

Case 1. Suppose $t_\ell \succ_s t \succ_s t_r$. Since the edge (s,t) is not stable while (s,t_ℓ) and (s,t_r) are stable edges, there is a stable matching M in H such that both s and t prefer their partners in M to each other (by Proposition 21). So $\mathsf{wt}_M(s,t) = -2$. This makes the edge (s,t) *slack* wrt to the popular matching M and its witness $y = 0$, i.e., $\mathsf{wt}_M(s,t) = -2 < 0 = y_s + y_t$.

Since $y = 0$ is a feasible solution to (LP2), $\mathsf{wt}_M(\tilde{N}) = \sum_{e \in \tilde{N}} \mathsf{wt}_M(e) < \sum_v y_v = 0$ (since $\mathsf{wt}_M(s,t) < y_s + y_t$). Thus $\Delta(N, M) < 0$, i.e., the stable matching M defeats N.

Case 2. Suppose $t \succ_s t_\ell$. That is, s prefers t to its most preferred stable partner t_ℓ in H. Consider the following two stable matchings in $H' = (A'_H \cup B'_H, E'_H)$:

$$M'_r = \{(a,b) : (a,b) \in M_r\} \cup \{(a', d(a)) : a \in A_H\}$$
$$M'_\ell = \{(a',b) : (a,b) \in M_\ell\} \cup \{(a, d(a)) : a \in A_H\}.$$

The vertex s' is matched to its top choice neighbor $d(s)$ in M'_r and it is matched to t_ℓ in M'_ℓ. Recall that $d(s) \succ_{s'} t \succ_{s'} t_\ell$. Since (s,t) is not a popular edge in H, the edge (s',t) is not stable in H'. We know that $(s', d(s))$ and (s', t_ℓ) are stable edges in H', hence there exists a stable matching M' in H' such that both s' and t prefer their respective partners in M' to each other (by Proposition 21). Observe that t's partner in M' has to be a *primed* neighbor (call it v') since t cannot prefer an *unprimed* neighbor to s'. So M' contains edges (s', u) and (v', t) where s' and t prefer their respective partners (u and v') to each other.

Let the stable matching M' in H' map to the popular matching M in H; let $y \in \{\pm 1\}^{n_H}$ be M's witness as described earlier. There are two subcases here.

- The vertex $u = d(s)$. So M' contains (s,b) (for some $b \in B_H$) and (v',t) where t prefers v' to s', i.e., t prefers v to s. The edges $(s,b), (v,t)$ are in M, where $\mathsf{wt}_M(s,t) \le 0$. We have $y_s = y_t = 1$ (by 1. and 2. stated earlier). Hence $\mathsf{wt}_M(s,t) \le 0 < 2 = y_s + y_t$.
- The vertex $u \ne d(s)$. So M' contains (s',u) and (v',t) where s prefers u to t and similarly, t prefers v to s. The edges $(s,u), (v,t)$ are in M and $\mathsf{wt}_M(s,t) = -2$. We have $y_s = -1$ and $y_t = 1$ (by 1. and 2. stated earlier). Hence $\mathsf{wt}_M(s,t) = -2 < 0 = y_s + y_t$.

So in both cases, the edge (s,t) is slack wrt M and its witness y. So complementary slackness (the same argument as given in case 1) implies that $\Delta(N, M) < 0$, i.e., the popular matching M defeats N.

Case 3. The last case is $t_r \succ_s t$. So s prefers its least preferred stable partner to t. If t also prefers its partner in M_r to s then M_r is a stable matching where both s and t prefer their respective partners to each other. This implies that M_r defeats N.

Else t prefers s to its partner in M_r, i.e., t prefers s to its most preferred stable partner. Observe that this is exactly the same as case 2 with the roles of s and t swapped. Thus an analogous argument shows that H has a popular matching that defeats N.

Proof of Lemma 14. Let M_c be a valid matching in G_c with a witness γ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm 2k\}$ for all $v \in C$. Recall that S (resp., U) is the set of stable (resp., unstable) vertices in G. We claim that all vertices in $S \cap C$ are matched in M_c and no vertex in $U \cap C$ is matched in M_c.

Consider (LP1) with M_c replacing M and $\tilde{E}_c = \tilde{E}_p \cap (C \times C)$ replacing $\tilde{E}_p$. The optimal value of this LP is 0 since there exists a dual feasible solution γ with $\sum_{u \in C} \gamma_u = 0$ (recall that γ obeys properties 1-3). Let N be a stable matching in G and let $N_c = N \cap (C \times C)$. If M_c leaves a vertex $v \in S \cap C$ unmatched then $\Delta(N_c, M_c) > 0$ (as shown in [18]), a contradiction to the optimal value of (LP1) being 0. Thus M_c matches all vertices in $S \cap C$. Since the self-loop $(u,u) \in \tilde{N}_c$ for any $u \in U \cap C$, the constraint $\gamma_u \ge \mathsf{wt}_{M_c}(u,u)$ is tight (by complementary slackness). Because $\mathsf{wt}_{M_c}(u,u) \in \{0, -1\}$ and γ_u is even, it follows that $\gamma_u = \mathsf{wt}_{M_c}(u,u) = 0$, i.e., u is left unmatched in M_c.

We need to show there is no blocking edge with respect to M'_c and this proof is similar to a proof in [24] on popular perfect matchings. Any dummy vertex $d_i(a)$ is matched either to its top choice neighbor a_{i-1} or to its second choice neighbor a_i; in the latter case, its top choice neighbor a_{i-1} is matched to a more preferred neighbor. Thus no blocking edge is incident to any dummy vertex. Let us now show that no blocking edge is incident to any other vertex in G'_c. Observe that $\tilde{M_c}$ is an optimal solution to (LP1), so for any $(p,q) \in M_c$, we have $\gamma_p + \gamma_q = \mathsf{wt}_{M_c}(p,q) = 0$ (by complementary slackness).

Let $a \in U \cap C_A$ and let $(a,b) \in E_c$. We have $\gamma_a = 0$ and $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a,b) \geq 0$. So $\gamma_b \geq 0$. If $\gamma_b = 0$ then $\mathsf{wt}_{M_c}(a,b) = 0$, i.e., $(z_0, \tilde{b}) \in M'_c$ for some neighbor z that b prefers to a. Else $\gamma_b > 0$ and so $(z_i, \tilde{b}) \in M'_c$ for some neighbor z with $-2i = \gamma_z = -\gamma_b < 0$, so $i > 0$, i.e., $\tilde{b}$ is matched to a neighbor in G'_c that it prefers to a_0. Hence $(a_0, \tilde{b})$ does not block M'_c.

Let us now show there is no blocking edge incident to a_ℓ, where $a \in S \cap C_A$ and $-k \leq \ell \leq k$. Suppose $\gamma_a = -2i$ and $(a,w) \in M_c$. Then $(a_i, \tilde{w}) \in M'_c$ and all of $a_{i+1}, \dots, a_k$ are matched to their respective top choice neighbors $d_{i+1}(a), \dots, d_k(a)$. Hence there is no blocking edge incident to a_j for $j \geq i+1$.

Let $(a,b) \in E_c$. If $b \in U$ then $\gamma_b = 0$ and $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a,b) \geq 0$. So $\gamma_a \geq 0$. If $\gamma_a = 0$ then $\mathsf{wt}_{M_c}(a,b) = 0$, i.e., $(a_0, \tilde{w}) \in M'_c$ for some neighbor w that a prefers to b. Else $\gamma_a > 0$ which implies that $i < 0$ and so $(a_0, d_0(a)) \in M_c$. In either case, a_0 prefers its partner in M'_c to $\tilde{b}$, so $(a_0, \tilde{b})$ does not block M'_c.

Let $b \in S$. Since $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a,b) \geq -2$, it follows that $\gamma_b \geq 2(i-1)$. Thus $(z_j, \tilde{b}) \in M'_c$ where $j \geq i-1$. Hence $\tilde{b}$ prefers its partner in M'_c to all a_j, where $j \leq i-2$. We now show that neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M'_c.

- If $j \geq i+1$ then $\tilde{b}$ prefers its partner z_j in M'_c to both a_i and a_{i-1}. Hence neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M'_c.
- If $j = i$ then $\gamma_a + \gamma_b = -2i + 2i = 0 \geq \mathsf{wt}_{M_c}(a,b)$. Thus either $(a_i, \tilde{b}) \in M'_c$ or one of $a_i, \tilde{b}$ prefers its partner in M'_c to the other. Hence neither $(a_i, \tilde{b})$ nor $(a_{i-1}, \tilde{b})$ blocks M'_c in this case as well.
- If $j = i-1$ then $\gamma_a + \gamma_b = -2i + 2(i-1) = -2 \geq \mathsf{wt}_{M_c}(a,b)$. So $\mathsf{wt}_{M_c}(a,b) = -2$, i.e., both a and b prefer their partners in M_c to each other. Hence $\tilde{b}$ prefers z_{i-1} to a_{i-1} and similarly, a_i prefers $\tilde{w}$ to $\tilde{b}$. Thus in this case also neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M'_c.

We prove the converse now. Let N be any stable matching in G and let $N_c = N \cap (C \times C)$. It is easy to check that $N'_c = \{(a_0, \tilde{b}) : (a,b) \in N_c\} \cup \{(a_i, d_{i+1}(a)) : a \in S \cap C_A \text{ and } i < 0\}$ $\cup \{(a_i, d_i(a)) : a \in S \cap C_A \text{ and } i > 0\}$ is a stable matching in G'_c. The set of vertices left unmatched in N'_c is $\{a_0, \tilde{b} : a, b \in U \cap C\}$. Hence the stable matching M'_c matches all vertices of G'_c except the vertices $a_0, \tilde{b}$, where $a, b \in U \cap C$.

In order to prove that M_c is valid in G_c, we define γ as follows:

- for every vertex $u \in U \cap C$, let $\gamma_u = 0$;
- for every edge $(p_i, \tilde{q}) \in M'_c$, let $\gamma_p = -2i$ and $\gamma_q = 2i$.

Since $-k \leq i \leq k$, it immediately follows that $\gamma_v \in \{0, \pm 2, \dots, \pm 2k\} \, \forall v \in C$. For any $u \in U \cap C$ (each such vertex is unmatched in M_c), we have $\gamma_u = 0 = \mathsf{wt}_{M_c}(u,u)$. We also have $\sum_{v \in C} \gamma_v = \sum_{(p,q) \in M_c} (\gamma_p + \gamma_q) = 0$.

Thus we are left to show the constraints $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a,b)$ for all $(a,b) \in E_c$. Then it will follow that properties 1-3 hold and thus M_c is valid in G_c with γ as a witness. Suppose $\gamma_a = -2i$ and $\gamma_b = 2j$. We need to show that $-2i + 2j \geq \mathsf{wt}_{M_c}(a,b)$ and this proof is similar to a proof in [23] on popular critical matchings. Let us consider the following 4 cases:

1. $j \geq i+1$: So $\gamma_a + \gamma_b \geq -2i + 2(i+1) = 2 \geq \mathsf{wt}_{M_c}(a,b)$ since $\mathsf{wt}_{M_c}(e) \in \{0, \pm 2\}$ for any $e \in E$.
2. $j = i$: Since the edge $(a_i, \tilde{b})$ does not block M'_c, either $(a_i, \tilde{b}) \in M'_c$ or one of $a_i, \tilde{b}$ is matched to a neighbor preferred to the other. Recall that the preference order of $\tilde{b}$ among level i neighbors in G'_c is exactly as per its preference order in G. Thus either $(a,b) \in M_c$ or one of a, b is matched in M_c to a neighbor preferred to the other. Hence $\gamma_a + \gamma_b = -2i + 2i = 0 \geq \mathsf{wt}_{M_c}(a,b)$.

3. $j = i - 1$: Observe that a has to be a stable vertex, otherwise $i = 0$ and the edge $(a_0, \tilde{b})$ would block M'_c. Since $j \geq -k$, we have $i = j + 1 \geq 1 - k$; so there is a vertex $d_i(a)$ which (as a_i's top choice) has to be matched in any stable matching in G'_c. Since $(a_i, \tilde{w}) \in M'_c$ for some $w \in B$, it follows that $(a_{i-1}, d_i(a)) \in M'_c$. So a_{i-1} is matched to its worst choice neighbor and because the edge $(a_{i-1}, \tilde{b})$ does not block M'_c, it follows that $(z_{i-1}, \tilde{b}) \in M'_c$ for some neighbor z that b prefers to a. The vertex $\tilde{b}$ prefers a_i to z_{i-1} since higher level neighbors are preferred to lower level neighbors. Since the edge $(a_i, \tilde{b})$ does not block M'_c, it follows that a prefers w to b. Thus both a and b prefer their respective partners in M_c to each other, so $\mathsf{wt}_{M_c}(a, b) = -2 = -2i + 2(i - 1) = \gamma_a + \gamma_b$.
4. $j \leq i - 2$: As argued in the above case, a has to be a stable vertex and $(a_{i-1}, d_i(a)) \in M'_c$. So a_{i-1} is matched to its worst choice neighbor. Either $\tilde{b}$ is unmatched or $(z_j, \tilde{b}) \in M'_c$ for some $j \leq i - 2$. In either case, M'_c has a blocking edge – a contradiction to its stability. Thus we cannot have $j \leq i - 2$. ◀

Proof of Lemma 15. The matching M_c has to match all vertices in $S \cap C$, otherwise we have $\Delta(N_c, M_c) > 0$ where N is any stable matching in G and $N_c = N \cap (C \times C)$, contradicting that there is a feasible solution γ to (LP2)[2] with $\sum_{v \in C} \gamma_v = 0$. Thus $\tilde{M}_c$ is feasible solution to (LP1); in fact, it is an optimal solution to (LP1) since $\mathsf{wt}_{M_c}(\tilde{M}_c) = \Delta(M_c, M_c) = 0$. If M_c leaves a vertex v unmatched then $(v, v) \in \tilde{M}_c$ and so by complementary slackness, we have $\gamma_v = \mathsf{wt}_{M_c}(v, v) = 0$. However all the γ-values are odd. Hence M_c matches all vertices in C.

We need to show that M''_c is stable in G''_c. As argued in the proof of Lemma 14, no blocking edge can be incident to any dummy vertex. Let us now show that there is no blocking edge incident to a_ℓ, where $a \in C_A$ and $\ell \geq -k$.

Suppose $(a, w) \in M_c$. Let $\gamma_a = -(2i - 1)$. Then $(a_i, \tilde{w}) \in M''_c$ and $(a_j, d_j(a)) \in M''_c$ for $j \geq i + 1$. Since a_j is matched to its top choice neighbor $d_j(a)$, there is no blocking edge incident to a_j for $j \geq i + 1$.

Let b be any neighbor of a in G_c, i.e., $(a, b) \in E_c$. Then $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a, b) \geq -2$, so $\gamma_b \geq 2i - 3 = 2(i - 1) - 1$. Thus $(z_j, \tilde{b}) \in M''_c$ where $j \geq i - 1$. Hence $\tilde{b}$ prefers its partner z_j to all a_ℓ, where $\ell \leq i - 2$. We now show that neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M''_c.

- If $j \geq i + 1$ then $\tilde{b}$ prefers its partner z_j in M''_c to both a_i and a_{i-1}. Hence neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M'_c.
- If $j = i$ then $\gamma_a + \gamma_b = -(2i - 1) + (2i - 1) = 0 \geq \mathsf{wt}_{M_c}(a, b)$. Thus either $(a_i, \tilde{b}) \in M''_c$ or one of $a_i, \tilde{b}$ prefers its partner in M''_c to the other. Hence neither $(a_i, \tilde{b})$ nor $(a_{i-1}, \tilde{b})$ blocks M''_c in this case.
- If $j = i - 1$ then $\mathsf{wt}_{M_c}(a, b) = -2$ and so both a and b prefer their partners in M_c to each other. So $\tilde{b}$ prefers z_{i-1} to a_{i-1} and similarly, a_i prefers $\tilde{w}$ to $\tilde{b}$. Thus in this case also neither $(a_{i-1}, \tilde{b})$ nor $(a_i, \tilde{b})$ blocks M''_c.

We will now prove the converse. We claim that M''_c is a perfect matching in G''_c. Let N be the max-size popular matching in G computed by the algorithm in [21]; N has a dual certificate in $\{0, \pm 1\}^n$ where every matched vertex has ± 1 in its coordinate. The matching $N_c = N \cap (C \times C)$ matches all the vertices in C (recall that $|C| \geq 2$).

We use N's dual certificate restricted to vertices in C (call this β, thus $\beta \in \{\pm 1\}^{|C|}$) to obtain a stable matching N''_c in G''_c. For $a \in C_A$, let $f(a) = (1 - \beta_a)/2$, so $f(a) = 0$ if $\beta_a = 1$, else $f(a) = 1$. Note that $f(a) = 1$ for every $a \in U \cap C_A$ (see [7, 21] for more details). Let $N''_c = \{(a_{f(a)}, \tilde{b}) : (a, b) \in N_c\} \cup \{(a_i, d_{i+1}(a)) : a \in C_A \text{ and } i < f(a)\} \cup \{(a_i, d_i(a)) : a \in S \cap C_A \text{ and } i > f(a)\}$. The stable matching N''_c matches all vertices in G''_c.

[2] This is the LP dual to (LP1) with M_c replacing M and $\tilde{E}_c = \tilde{E}_p \cap (C \times C)$ replacing $\tilde{E}_p$.

Hence the stable matching M_c'' is also a perfect matching in G_c''. Thus M_c matches all vertices in C. In order to prove that M_c is a valid matching in G_c, we define γ as follows:

- for every edge $(p_i, \tilde{q}) \in M_c''$, let $\gamma_p = -(2i-1)$ and $\gamma_q = 2i - 1$.

Since $-k \leq i \leq k+1$, we have $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k+1)\} \, \forall v \in C$. We also have $\sum_{v \in C} \gamma_v = \sum_{(p,q) \in M_c} (\gamma_p + \gamma_q) = 0$.

Furthermore, for any $a \in U \cap C_A$, we have $(a_i, \tilde{w}) \in M_c''$ where $-k \leq i \leq 1$ for some neighbor w; thus $\gamma_a = -(2i-1) \geq -1$. Similarly, for any $b \in U \cap C_B$, we have $(z_j, \tilde{b}) \in M_c''$ where $0 \leq j \leq k+1$ for some neighbor z; thus $\gamma_b = 2j - 1 \geq -1$. Hence for any $u \in U \cap C$, we have $\gamma_u \geq -1 = \mathsf{wt}_{M_c}(u, u)$.

Thus we are left to show the constraints $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a, b)$ for all $(a, b) \in E_c$. Then it will follow that properties 1-3 hold and thus M_c is valid in G_c with γ as a witness. Suppose $\gamma_a = -(2i-1)$ and $\gamma_b = 2j - 1$. As done in the proof of Lemma 14, let us consider the following 4 cases:

1. $j \geq i+1$: So $\gamma_a + \gamma_b \geq 2 \geq \mathsf{wt}_{M_c}(a, b)$ since $\mathsf{wt}_{M_c}(e) \in \{0, \pm 2\}$ for any $e \in E$.
2. $j = i$: Since the edge $(a_i, \tilde{b})$ does not block M_c'', either $(a_i, \tilde{b}) \in M_c''$ or one of $a_i, \tilde{b}$ is matched to a neighbor preferred to the other. Thus either $(a, b) \in M_c$ or one of a, b is matched in M_c to a neighbor preferred to the other. So $\gamma_a + \gamma_b = -(2i-1) + 2i - 1 = 0 \geq \mathsf{wt}_{M_c}(a, b)$.
3. $j = i - 1$: So $(a_i, \tilde{w})$ and $(z_{i-1}, \tilde{b})$ are in M_c''. The vertex $\tilde{b}$ prefers a_i to z_{i-1} because higher level neighbors are preferred to lower level neighbors. Since the edge $(a_i, \tilde{b})$ does not block M_c'', it follows that a_i prefers $\tilde{w}$ to $\tilde{b}$. Observe that $(a_{i-1}, d(a_i)) \in M_c''$, thus a_{i-1} is matched to its worst choice neighbor $d(a_i)$. Since the edge $(a_{i-1}, \tilde{b})$ does not block M_c'', it follows that $\tilde{b}$ prefers z_{i-1} to a_{i-1}. Thus both a and b prefer their respective partners in M_c to each other, so $\mathsf{wt}_{M_c}(a, b) = -2 = -(2i-1) + 2(i-1) - 1 = \gamma_a + \gamma_b$.
4. $j \leq i - 2$: We have $(z_j, \tilde{b}) \in M_c''$ for some $j \leq i - 2$. As argued in the previous case, the edge $(a_{i-1}, d(a_i)) \in M_c''$. This means that M_c'' has a blocking edge, a contradiction to its stability. Hence this case does not occur. ◀