Matchings, Relaxed Popularity, and Optimality*

Telikepalli Kavitha

Tata Institute of Fundamental Research, Mumbai, India
kavitha@tifr.res.in

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 to N is at least the number that prefer N to M; 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 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 a 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 it is NP-complete to decide if there exists a popular matching that is more popular than a given matching. Interestingly, there exists a set 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 this set.

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 (i) either a is unmatched or prefers b to its partner in M and (ii) either b is unmatched or prefers a to its partner in M. The existence of stable matchings in a marriage instance and the Gale-Shapley algorithm [17] to find one are classic 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 [7,32]. Stability is a rather strict notion—all stable matchings match the same subset of vertices [18] and the size of a stable matching might be only half the size of a maximum matching. In applications such as matching students to advisers, 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 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 $\Delta(M, N) = \phi(M, N) - \phi(N, M)$.

* A preliminary version of this paper appeared in STACS 2022 as "Fairly popular matchings and optimality" [28].
Funding. Supported by the Department of Atomic Energy, Government of India, under project no. RTI4001.
Acknowledgments. Thanks to Yuri Faenza and Jaikumar Radhakrishnan for useful discussions. Thanks to the reviewers of this version and also of the conference version of the paper for their helpful comments and suggestions.

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 [19] introduced the notion of popularity in 1975 where he observed that every stable matching is popular. In fact, stable matchings are min-size popular matchings [21]. 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 [9] 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 [10,21,24]. Many natural optimization problems in popular matchings such as the min-cost popular matching problem are NP-hard [13]; moreover, this problem is NP-hard to approximate to any multiplicative factor. Though relaxing stability to popularity promises matchings with improved optimality with respect to cost, finding these matchings is hard.

The extension complexity of the popular matching polytope of G is $2^{\Omega(m/\log m)}$ [12]. 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 [11] (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 "mild 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) We can efficiently optimize over matchings M that satisfy $\Delta(M,S) \geq 0$ for all $S \in \mathcal{S}$.
(c) For any matching $N \notin \mathcal{S}$, there is at least one matching $S \in \mathcal{S}$ such that $\Delta(N,S) < 0$.

Remark 1. Note that property (c) is independent of property (a); the latter says that every matching $N \notin \mathcal{S}$ has to fail our *test of mild popularity* (this test is yet to be defined) while the former says that any matching $N \notin \mathcal{S}$ has to be defeated by a matching in $\mathcal{S}$. Property (c) will ensure that our matching M (so $\Delta(M,S) \geq 0$ for all $S \in \mathcal{S}$) is in $\mathcal{S}$. Without property (c), we may end up with a matching that does not pass our test of mild popularity.

Thus we should define our test of mild popularity such that any matching M that satisfies $\Delta(M,S) \geq 0$ for all $S \in \mathcal{S}$ will pass this test. For example, if $\mathcal{S} = \{\text{popular matchings}\}$, then it is not the case that every matching undefeated by all popular matchings has to be popular—Section 1.2 has such an example. Thus property (c) does not hold if we set *popularity* as our criterion of mild popularity.

The unpopularity of a matching M is typically measured by its *unpopularity factor* [31], which is defined as $u(M) = \max_{N \neq M} \phi(N,M)/\phi(M,N)$. A matching M is popular if and only if $u(M) \leq 1$.

Suppose we define a matching M to be very unpopular if $u(M) = \infty$, in other words, let $\mathcal{S} = \{\text{Pareto optimal matchings}\}$.[1] Observe that any matching M undefeated by all Pareto optimal matchings has to be Pareto optimal, in fact, M has to be popular. So it is NP-hard to find a min-cost matching M such that $\Delta(M,S) \geq 0$ for all $S \in \{\text{Pareto optimal matchings}\}$. Hence property (b) does not hold if we set *Pareto optimality* as our criterion of mild popularity.

[1] A matching M is Pareto optimal if there is no matching N such that $\phi(N,M) > 0$ and $\phi(M,N) = 0$.

1.1 Our main results

A *mixed* matching $\varPi$ is a probability distribution or a lottery over matchings, so $\varPi = \{(M_0, p_0), \ldots, (M_k, p_k)\}$ 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 [30]; the mixed matching $\varPi$ is popular if $\Delta(\varPi, N) = \sum_{i=0}^{k} p_i \cdot \Delta(M_i, N) \geq 0$ for all matchings N. We will use popular mixed matchings to define a natural relaxation of popularity.

The matchings $M_0, \ldots, M_k$ are said to be in the support of $\varPi = \{(M_0, p_0), \ldots, (M_k, p_k)\}$. Let us call a matching M *supporting* if there exists a popular mixed matching $\varPi$ 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. So our set $\mathcal{S}$ of relevant matchings will be the set of all supporting matchings, i.e., $\mathcal{S} = \{S : S \text{ is a supporting matching}\}$. It is easy to see that the set $\mathcal{S}$ is sandwiched between the set of popular matchings and the set of Pareto optimal matchings.

We are ready to define *fairly popular* matchings now.

Definition 2. *A matching M is fairly popular if $\Delta(M, S) \geq 0$ for all $S \in \mathcal{S}$ where $\mathcal{S}$ is the set of supporting matchings.*

For any matching N that defeats a fairly popular matching M, it is the case that even with the help of other matchings, N cannot form a popular mixture. Thus it is natural to regard a *non-supporting* matching N as being "very unpopular". Hence elections against non-supporting matchings will not be relevant. Intriguingly, waiving the constraints $\Delta(M, N) \geq 0$ for non-supporting matchings N makes the resulting polytope easy to describe.

Theorem 1. *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 2). Any point $\vec{x} \in \mathbb{R}^m_{\geq 0}$ such that $\sum_{e \in \delta(v)} x_e \leq 1$ for each vertex v is a *fractional* matching and $\vec{x}$ is equivalent to a mixed matching (Birkhoff-von Neumann theorem [8]). A fractional matching $\vec{x}$ is popular if $\varPi$ is a popular mixed matching, where $\varPi$ is any mixed matching that corresponds to $\vec{x}$ (see [30]). The following terms will be useful to us.

- An edge e is a *popular fractional* edge if there exists a popular fractional matching $\vec{x}$ with $x_e > 0$.
- A vertex v is *stable* if v is matched in any stable matching in G. All stable matchings match the same subset of vertices [18], so unstable vertices are left unmatched in every stable matching.

Theorem 2. *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., $\Delta(\varPi, M) \leq 0$ for all popular mixed matchings $\varPi$.*[2]
3. *M matches all stable vertices and $M \subseteq E_p$, where E_p is the set of popular fractional edges.*

[2] Equivalently, $\Delta(\varPi, M) = 0$ since $\Delta(\varPi, M) \geq 0$ for all matchings M because $\varPi$ is a popular mixed matching.

Remark 2. Theorem 2 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). Thus by Theorem 1 and Theorem 2, the set $\mathcal{S} = \{$supporting matchings$\}$ satisfies properties (b) and (c) stated earlier. Hence every fairly popular matching is also supporting since no supporting matching defeats a fairly popular matching (by definition).

1.2 Our other results

Consider the following instance from [22] where $A = \{a_0, a_1, a_2\}, B = \{b_0, b_1\}$, and vertex preferences are as follows:

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

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. This instance has only one popular matching $P = \{(a_0, b_0), (a_1, b_1)\}$. Observe that P is more popular than $N = \{(a_0, b_0), (a_2, b_1)\}$ and N is more popular than $M = \{(a_0, b_1), (a_1, b_0)\}$, but P is *not* more popular than M. Thus M is undefeated by the only popular matching P. So it is not the case that every unpopular matching has to be defeated by some popular matching.

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 note that N is not a supporting matching (since N leaves the stable vertex a_1 unmatched).

Suppose we had defined our set of relevant matchings to be the set of matchings undefeated by popular matchings. This is a superset of our set $\mathcal{S}$ which—by Theorem 2—is the set of matchings undefeated by a larger set: the set of popular mixed matchings. To be undefeated by popular matchings is a natural threshold for mild popularity as any matching defeated by a popular matching can be considered to be *very unpopular*.

Before we check whether such a set of relevant matchings obeys the desired properties (b)-(c) stated earlier, let us ask how easy it is to test membership in this set. That is, given a matching N, is it easy to determine if there exists a popular matching that defeats N? Interestingly, we can show a "compactness" result. Note that G may have more than 2^n popular matchings [36].

Proposition 1. *There is a set of at most m popular matchings in G such that any matching defeated by some popular matching in G has to be defeated by one of these m popular matchings.*

However, deciding if a given matching is undefeated by all popular matchings is coNP-complete.

Theorem 3. *Given a marriage instance $G = (A \cup B, E)$ and a matching N in G, it is NP-complete to decide if there exists any popular matching that is more popular than N.*

So if we had defined our set $\mathcal{S}$ of relevant matchings to be those undefeated by popular matchings, then it would have been coNP-hard to identify which matchings are in $\mathcal{S}$ (by Theorem 3). By letting $\mathcal{S} = \{$matchings undefeated by popular mixed matchings$\}$, we have a natural strengthening of the above notion of mild popularity. Moreover, as shown in Theorem 2, the matchings in our set $\mathcal{S}$ satisfy another natural and our original notion of mild popularity (property 1 of Theorem 2) and have a simple and clean combinatorial characterization (property 3 of Theorem 2).

1.3 Related results

The min-cost stable matching problem is very well-studied with several polynomial time algorithms [14,15,16,23,37] to solve this problem; furthermore, the stable matching polytope has a simple and elegant linear size formulation in $\mathbb{R}^m$ [33,35]. In contrast to this, as mentioned earlier, the extension complexity of the popular matching polytope of G is $2^{\Omega(m/\log m)}$ [12]. It is known that the popular fractional matching polytope of G is half-integral [22].

A min-cost popular matching in G can be computed in $O^*(2^{n/4})$ time [27]. The intractability of the min-cost popular matching problem has motivated relaxations such as *quasi-popularity* [12] and *semi-popularity* [27]. 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 [12]. 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 [27] to find an *almost* semi-popular matching whose cost is at most twice the cost of a min-cost popular matching.

Popular mixed matchings were introduced in [30] in the setting of *one-sided* popular matchings in a bipartite instance $G = (A \cup B, E)$. So it is only vertices in A that have preferences—popular matchings need not always exist in such a setting. It was shown in [30] that popular mixed matchings always exist and such a mixed matching can be computed in polynomial time.

1.4 Our techniques

Our main novelty is in our characterization of supporting matchings—this leads to a characterization of fairly popular matchings. The characterization of supporting matchings (given in Section 2) uses the half-integrality of the popular fractional matching polytope in a marriage instance [22] along with Hall's marriage theorem on perfect matchings in bipartite graphs. The main technical lemma here is based on the existence of certain helpful stable matchings as shown in [20].

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 whose edge set is restricted to the set E_p of popular fractional edges. Every matching M_c in the decomposition $M = \cup_C M_c$ has a certain *witness* that is obtained via LP duality. The LP-machinery for popular matchings was introduced in [30] and used in [22,25] to study popular fractional matchings.

We define two *colorful multigraphs* G'_c and G''_c where each edge is assigned a color—these multigraphs are inspired by instances from [26,29] that solve variants of the popular matching problem by modeling them as stable matching problems in appropriate graphs. In particular, the min-cost popular *maximum* matching problem was studied in [29]. It was shown in [24] that there always exists a maximum matching that is popular within the set of maximum matchings and a polynomial time algorithm to find a *min-cost* such matching was given in [29] by modeling it as a min-cost stable matching problem in an appropriate multigraph.

Our algorithm follows the same outline as in [29]. However there is no single graph that we can construct such that every fairly popular matching M is a stable matching in the new graph. We use *witnesses* (mentioned earlier) for matchings M_c in $M = \cup_C M_c$ to show a surjective mapping from the union of sets of stable matchings in the colorful multigraphs 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] (stated in Section 3.3) to formulate the convex hull of $\mathcal{S}'_c \cup \mathcal{S}''_c$ leads to

Theorem 1 (proved in Section 3). Thus, unlike the popular matching polytope, the fairly popular matching polytope $\mathcal{F}$ has a compact extended formulation.

Our NP-hardness proof is given in Section 4. This is based on the NP-hardness (from [13]) of deciding if there exists a popular matching that contains two forced edges.

2 A Characterization of Supporting Matchings

We prove Theorem 2 in this section. Before we characterize supporting matchings, it will be useful to recall some properties of popular fractional matchings in a marriage instance $G = (A \cup B, E)$.

Fractional matchings. A fractional matching $\vec{x}$ in G is a convex combination of matchings (by Birkhoff-von Neumann theorem [8]). Recall that $\vec{x}$ is popular if Π is a popular mixed matching, where Π is any mixed matching that is equivalent to $\vec{x}$.

Alternatively, as shown in [30], $\vec{x}$ is popular if $\Delta(\vec{x}, M) \geq 0$ for all matchings M. In order to define $\Delta(\vec{x}, M)$, we need to first define the function $\mathsf{vote}_u(v, M)$.

– For any vertex u and a neighbor v of u, the value $\mathsf{vote}_u(v, M)$ is 1 if u prefers v to its assignment in M, it is -1 if u prefers its assignment in M to v, and it is 0 otherwise (i.e., $v = M(u)$).

It will be convenient to assume that $\vec{x}$ fully matches u, so let us set $x_{(u,u)} = 1 - \sum_{e \in \delta(u)} x_e$ where $\delta(u)$ is the set of edges incident to u in G. Thus $\vec{x}$ matches u to itself with fractional weight $x_{(u,u)}$ and u considers being matched to itself as its worst choice (i.e., equivalent to being left unmatched).

$$\text{Let } \mathsf{vote}_u(\vec{x}, M) = \sum_{(u,v) \in \delta(u) \cup \{(u,u)\}} x_{(u,v)} \cdot \mathsf{vote}_u(v, M).$$

Let $\Delta(\vec{x}, M) = \sum_{u \in A \cup B} \mathsf{vote}_u(\vec{x}, M)$. Recall that $\vec{x}$ is popular if $\Delta(\vec{x}, M) \geq 0$ for all matchings M. The popular fractional matching polytope of G is the convex hull of all popular fractional matchings $\vec{x}$ in G. It was shown in [22] that the popular fractional matching polytope of G is half-integral. This proof of half-integrality uses the graph $H = (A_H \cup B_H, E_H)$ defined below.

The graph H. The graph H can be regarded as consisting of *two* copies of $G = (A \cup B, E)$ (see Fig. 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.

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}.

Let N be any matching in G. Corresponding to N, there is a perfect matching N' in H defined as $\{(a_0, b_0), (a_1, b_1) : (a, b) \in N\} \cup \{(u_0, u_1) : u \text{ is unmatched in } N\}$. If N is a stable matching in G, then it is easy to see that N' is a stable matching in H. Thus H admits a *perfect* stable matching, i.e., one that matches all vertices. It was shown in [22, 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.

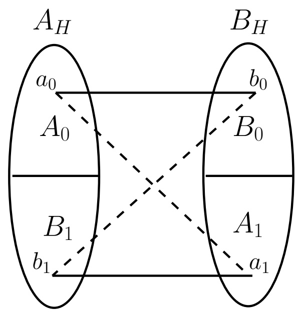

Fig. 1. The vertex set of H has 2 copies u_0 and u_1 of every vertex u in G and 2 copies $(a_0, b_0), (a_1, b_1)$ of each edge (a, b) in G along with the edges (u_0, u_1) for all u (these are the dashed edges).

The above map from matchings in G to matchings in H extends to fractional matchings. So for any fractional matching $\vec{x}$ in G, there is a corresponding fractional matching $\vec{x}'$ in H where

$$\begin{aligned} x'_{(u_0,v_0)} &= x'_{(u_1,v_1)} = x_{(u,v)} && \forall\, (u,v) \in E; \\ x'_{(u_0,u_1)} &= x_{(u,u)} = 1 - \sum_{e \in \delta(u)} x_e && \forall\, u \in A \cup B. \end{aligned}$$

The following claim will be useful. Note that an edge e is said to be a *popular edge* if there is a popular matching containing e.

Claim 1 *The edge $(a, b) \in E_p$ if and only if (a_0, b_0) and (a_1, b_1) are popular edges in H.*

Proof. If $\vec{x}$ is a popular fractional matching in G then observe that $\vec{x}'$ is a popular fractional matching in H. This is because for any matching N in H, $\Delta(\vec{x}', N) = \Delta(\vec{x}, N_0) + \Delta(\vec{x}, N_1)$ where for $i \in \{0, 1\}$, N_i is the subset of N in the subgraph of H induced on subscript i vertices. Hence if $(a, b) \in E_p$, i.e., if (a, b) is a popular fractional edge in G, then (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 in H.

Conversely, suppose (a_0, b_0) is a popular edge in H. Then there is a popular matching P in H containing the edge (a_0, b_0). Note that P is a perfect matching and let $\vec{p}$ be its edge incidence vector. Define the fractional matching $\vec{r}$ in G as follows: $r_{(u,v)} = (p_{(u_0,v_0)} + p_{(u_1,v_1)})/2$ for any $(u, v) \in E$ and $r_{(u,u)} = p_{(u_0,u_1)}$ for any $u \in A \cup B$. For any matching N in G, observe that $\Delta(\vec{r}, N) = \Delta(P, N')/2$. Since P is popular in H, we have $\Delta(P, N') \geq 0$ and thus $\Delta(\vec{r}, N) \geq 0$. Hence $\vec{r}$ is a popular fractional matching in G, so $(a, b) \in E_p$. □

Remark 3. Note that the edge (u_0, u_1) is popular in H if and only if u is an unstable vertex in G. For any stable matching S in G, recall that the matching S' is stable (hence, popular) in H and it contains the edges (u_0, u_1) for all unstable vertices u; moreover, $(v, v) \notin E_p$ for any stable vertex v [22, Footnote 2].

2.1 Proof of Theorem 2

We need to show the following three statements are equivalent in $G = (A \cup B, E)$.

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., $\Delta(\Pi', M) > 0$.

Since both Π and Π' are popular mixed matchings, we have $\Delta(\Pi', \Pi) = \sum_j p_j \cdot \Delta(\Pi', M_j) = 0$. Because $\Delta(\Pi', M_i) > 0$ 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 $1 \Rightarrow 2$.

Proof of 2⇒3. This part needs the following technical lemma. Call an edge e *unpopular* if there exists no popular matching that contains e.

Lemma 1. *Any matching in H that contains an unpopular edge is defeated by some popular matching in H.*

For now, we will assume Lemma 1 and finish the proof of Theorem 2. The proof of Lemma 1 is given in Section 2.2.

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 *unpopular* edge (by Claim 1). Thus some popular matching P in H defeats M' (by Lemma 1).

Let $\vec{p}$ be the edge incidence vector of P. Define the fractional matching $\vec{r}$ in G as follows: $r_{(a,b)} = (p_{(a_0,b_0)} + p_{(a_1,b_1)})/2$ for any $(a, b) \in E$ and $r_{(u,u)} = p_{(u_0,u_1)}$ for any $u \in A \cup B$. We have $\Delta(\vec{r}, N) = \Delta(P, N')/2$ for any matching N in G, so $\vec{r}$ is a popular fractional matching in G. Furthermore, we have $\Delta(\vec{r}, M) > 0$ since $\Delta(P, M') > 0$. The fractional matching $\vec{r}$ is equivalent to a mixed matching Π, note that Π is popular since $\vec{r}$ is popular. Thus there is a popular mixed matching Π more popular than M, a contradiction to M satisfying property 2. Thus $2 \Rightarrow 3$.

Proof of 3⇒1. Every vertex left unmatched in M is unstable in G, so there is a popular matching S' in H that contains all the edges (u_0, u_1) where u is unmatched in M (see Remark 3). Each edge $e = (a, b) \in M$ belongs to E_p (because $M \subseteq E_p$). So 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 (by Claim 1).

Let $M = \{e_1, \ldots, e_\ell\}$. Consider the 2ℓ matchings $M_{e_1}^0, \ldots, M_{e_\ell}^0, M_{e_1}^1, \ldots, M_{e_\ell}^1$ analogous to the matchings M_e^0 and M_e^1, defined above in the graph H. Let H' be the graph whose edge set is the multiset of edges present in these 2ℓ popular matchings and the popular matching S'. So 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 $(2\ell+1)$-regular since each of these $2\ell+1$ 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 $M' = \{(a_0, b_0), (a_1, b_1) : (a, b) \in M\} \cup \{(u_0, u_1) : u$ 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 2ℓ-regular. It follows from Hall's marriage theorem that H'' can be decomposed into 2ℓ perfect matchings $N_1', \ldots, N_{2\ell}'$ [5, Exercise 11.12]. Thus we have:

$$I_{M'} + I_{N_1'} + \cdots + I_{N_{2\ell}'} \;=\; I_{M_{e_1}^0} + \cdots + I_{M_{e_\ell}^0} + I_{M_{e_1}^1} + \cdots + I_{M_{e_\ell}^1} + I_{S'},$$

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

The $2\ell+1$ matchings $M_{e_1}^0, \ldots, M_{e_\ell}^1, S'$ (on the right hand side above) are popular in H. Hence the fractional matching $\vec{q} = (I_{M_{e_1}^0} + \cdots + I_{S'})/(2\ell+1)$, which is a convex combination of these matchings, is popular in H. Note that $\vec{q}$ can also be written as $(I_{M'} + I_{N_1'} + \cdots + I_{N_{2\ell}'})/(2\ell+1)$, where $M', N_1', \ldots, N_{2\ell}'$ are the matchings on the left hand side of this equation.

Consider the fractional matching $\vec{r}$ in G defined as $r_{(a,b)} = (q_{(a_0,b_0)} + q_{(a_1,b_1)})/2$ for any $(a,b) \in E$ and $r_{(u,u)} = q_{(u_0,u_1)}$ for any $u \in A \cup B$. As seen in the proof of Claim 1, the popularity of $\vec{q}$ in H implies the popularity of $\vec{r}$ in G. Consider the mixed matching $\Pi = \{(M, \frac{1}{2\ell+1}), \ldots\}$ that is equivalent to $\vec{r}$. Since $\vec{r}$ is popular, Π is popular and Π has support on M. Thus M is a supporting matching. Hence 3 ⇒ 1. ◻

2.2 Proof of Lemma 1

We need to show that any matching in H that contains an unpopular edge is defeated by some popular matching in H. Before we formally prove this lemma, we give a high level intuition of its proof. Any popular matching M (augmented with self-loops at unmatched vertices) is a max-weight perfect matching as per a certain edge weight function wt_M defined below. Thus M can be realized as an optimal solution to a linear program (see (LP1) below).

An optimal solution to the dual LP is a *dual certificate* for M. As proved in Theorem 4, a popular matching M in a marriage instance on k vertices admits a dual certificate in $\{0, \pm 1\}^k$. Popular matchings in the instance H are perfect matchings and they admit dual certificates in $\{\pm 1\}^{2n}$ [22] (recall that H has $2n$ vertices). This simpler dual certificate allows us to realize any popular matching in H as a stable matching in an auxiliary marriage instance H^* [10]. An unpopular edge in H becomes an *unstable* edge in H^*, i.e., no stable matching contains it.

Now we can use the machinery of stable matchings. It was shown in [20] that if (s,t) is an unstable edge in a marriage instance such that (s,t_0) and (s,t_1) are stable edges for some neighbors t_0, t_1 of s where $t_1 \succ_s t \succ_s t_0$ then there exists a stable matching where both s and t prefer their partners to each other. This stable matching in H^* will lead to our desired popular matching in H.

We now formally discuss the preliminaries that will be used in our proof. Let $\tilde{H} = (A_H \cup B_H, \tilde{E}_H)$ be the graph H augmented with self-loops at all vertices. So each vertex u regards itself as its last choice neighbor and any matching M in H becomes a perfect matching $\tilde{M}$ in $\tilde{H}$ by augmenting M with self-loops at vertices left unmatched in M. For any matching M, the following edge weight function wt_M can be defined. For each edge $(a,b) \in E_H$:

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

Observe that $\mathsf{wt}_M(u,v) = \mathsf{vote}_u(v,M) + \mathsf{vote}_v(u,M)$ for any edge $(u,v) \in E_H$, where the function $\mathsf{vote}_u(v,M)$ was defined earlier in Section 2. For each vertex u, let $\mathsf{wt}_M(u,u) = 0$ if u is left unmatched in M, else $\mathsf{wt}_M(u,u) = -1$. So for any $u \in A_H \cup B_H$, we have $\mathsf{wt}_M(u,u) = \mathsf{vote}_u(u,M)$.

Let N be any matching in H. We have:

$$\mathsf{wt}_M(\tilde{N}) = \sum_{u \in A_H \cup B_H} \mathsf{vote}_u(\tilde{N}(u), \tilde{M}(u)) = \phi(N,M) - \phi(M,N) = \Delta(N,M).$$

So M is popular in H if and only if $\mathsf{wt}_M(\tilde{N}) \le 0$ for all matchings N in H. Consider the following linear program where $\delta_H(u)$ is the set of edges incident to u in H.

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

subject to

$$\sum_{e \in \delta_H(u) \cup \{(u,u)\}} x_e = 1 \quad \forall\, u \in A_H \cup B_H \qquad \text{and} \qquad x_e \geq 0 \quad \forall\, e \in \tilde{E}_H.$$

The constraint matrix of (LP1) is totally unimodular. This is because H is a bipartite graph and adding self-loops to this graph preserves the total unimodularity of the constraint matrix. So this LP computes a max-weight perfect matching in $\tilde{H}$ with respect to the edge weight function wt_M. Thus matching M is popular in H if and only if the optimal value of (LP1) is at most 0. In fact, the optimal value is exactly 0 since $\tilde{M}$ is a perfect matching in $\tilde{H}$ and $\mathsf{wt}_M(\tilde{M}) = 0$ because $\mathsf{wt}_M(e) = 0$ for each edge/self-loop e in $\tilde{M}$.

The linear program (LP2) is the dual LP. By LP duality, M is popular in H if and only if there exists a dual feasible solution $\vec{y} \in \mathbb{R}^{2n}$ such that $\sum_{u \in A_H \cup B_H} y_u = 0$ (recall that $|A_H \cup B_H| = 2n$).

$$\text{minimize} \sum_{u \in A_H \cup B_H} y_u \tag{LP2}$$

subject to

$$y_a + y_b \;\geq\; \mathsf{wt}_M(a,b) \quad \forall\,(a,b) \in E_H \qquad \text{and} \qquad y_u \geq \mathsf{wt}_M(u,u) \quad \forall\, u \in A_H \cup B_H.$$

Theorem 4 ([25]). *A matching M in $H = (A_H \cup B_H, E_H)$ is popular if and only if there exists $\vec{y} \in \{0, \pm 1\}^{2n}$ such that $\sum_{u \in A_H \cup B_H} y_u = 0$ along with $y_a + y_b \geq \mathsf{wt}_M(a,b)$ for all $(a,b) \in E_H$ and $y_u \geq \mathsf{wt}_M(u,u)$ for all $u \in A_H \cup B_H$.*

Proof. The constraint matrix of (LP2) is totally unimodular. So (LP2) admits an optimal solution that is integral. Let $\vec{y}$ be an integral optimal solution of (LP2). Thus $\vec{y} \in \mathbb{Z}^{2n}$.

We need to show that $\vec{y} \in \{0, \pm 1\}^{2n}$. We have $y_u \geq \mathsf{wt}_M(u,u) \geq -1$ for all $u \in A_H \cup B_H$. Since $\tilde{M}$ is an optimal solution to (LP1), complementary slackness implies that $y_u + y_v = \mathsf{wt}_M(u,v) = 0$ for every $(u,v) \in \tilde{M}$. Thus $y_u = -y_v \leq 1$ for every vertex u matched to a non-trivial neighbor v in $\tilde{M}$. Regarding any vertex u such that $(u,u) \in \tilde{M}$, we again have by complementary slackness $y_u = \mathsf{wt}_M(u,u) = 0$. Hence $\vec{y} \in \{0, \pm 1\}^{2n}$. ☐

We will call a vector $\vec{y}$, as given in Theorem 4, a *dual certificate* for popular matching M. It was shown in [22, Lemma 2] that every popular matching in H has a dual certificate $\vec{y} \in \{\pm 1\}^{2n}$ (this uses the fact that H admits a perfect stable matching).

An auxiliary instance. Since every popular matching in H is perfect, there is a surjective map (as shown in [10]) 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 sets A'_H and B'_H are defined below.

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

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

– Then the preference order of a in H^* is $b_1 \succ \cdots \succ b_r \succ d(a)$.
– And the preference order of a' in H^* is $d(a) \succ b_1 \succ \cdots \succ b_r$.

Let $b \in B_H$. Suppose b's preference order in H is $a_1 \succ \cdots \succ a_k$.

– Then the preference order of b 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.

Recall that any popular matching M in H has a dual certificate $\vec{y} \in \{\pm 1\}^{2n}$.

$$\text{Let } M' = \bigcup_{\substack{a \in A_H \\ y_a = 1}} \{(a,b),(a',d(a)) : (a,b) \in M\} \bigcup_{\substack{a \in A_H \\ y_a = -1}} \{(a',b),(a,d(a)) : (a,b) \in M\}.$$

It was shown in [10, Lemma 5] that M' is a stable matching in H^*. Conversely, let M' be any stable matching in H^*. Then M' projects to the matching $M = \{(a,b) : (a,b) \text{ or } (a',b) \text{ is in } M'\}$ in H. The popularity of M in H can be proved via the following vector $\vec{y}$:

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 $y_b = -1$.

Observe that $y_a + y_b = 0$ for each edge $(a,b) \in M$. Since M is a perfect matching, we have $\sum_{u \in A_H \cup B_H} y_u = 0$. We refer to [22, Section 3] for the details that $\vec{y}$ is a feasible solution to (LP2). Since $\mathsf{wt}_M(\tilde{M}) = \sum_{u \in A_H \cup B_H} y_u = 0$, it follows that the incidence vector of $\tilde{M}$ is an optimal solution to (LP1) and $\vec{y}$ is an optimal solution to (LP2). Thus M is a popular matching in H with $\vec{y}$ as a dual certificate.

We are now ready to prove Lemma 1. Let (s,t) be an *unpopular* edge in H. For any matching N that contains (s,t), we will show a popular matching more popular than N. The following result on stable matchings in a marriage instance will be useful to us. Call an edge e *stable* if there is a stable matching in H that contains e.

Proposition 2. *[20, 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.*

We will consider three cases based on the position of t in s's preference order on its neighbors in H. In each case we will use Proposition 2 to construct a desired popular matching in H.

Proof of Lemma 1. Let N be a matching in H that contains an unpopular edge (s,t). 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 2). So $\mathsf{wt}_M(s,t) = -2$. Observe that the edge (s,t) is *slack* with respect to the popular matching M and its dual certificate $\vec{y} = \vec{0}$.[3] That is:

$$\mathsf{wt}_M(s,t) = -2 \; < \; 0 = y_s + y_t.$$

[3] Since matching M is stable, we have $\mathsf{wt}_M(e) \le 0$ for all edges e; thus the vector $\vec{0}$ is a dual certificate for M.

So we have $\mathsf{wt}_M(\tilde{N}) = \sum_{e\in\tilde{N}} \mathsf{wt}_M(e) < \sum_u y_u = 0$ since $\mathsf{wt}_M(s,t) < y_s + y_t$ and $\mathsf{wt}_M(a,b) \le y_a + y_b$ for all edges (a,b) (since $\vec{y}$ is a feasible solution to (LP2)). 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 in the graph H^*, we have $d(s) \succ_{s'} t \succ_{s'} t_\ell$. We know that $(s', d(s))$ and (s', t_ℓ) are stable edges in H^* since $(s', d(s)) \in M'_r$ and $(s', t_\ell) \in M'_\ell$. However, (s', t) is not a stable edge in H^* since (s,t) is not a popular edge 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 2).

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.

The stable matching M' in H^* projects to a popular matching M in H; let $\vec{y} \in \{\pm 1\}^{2n}$ be M's witness as described in points 1 and 2 just before the proof of Lemma 1. There are two subcases.

- 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 the definition of $\vec{y}$. Hence $\mathsf{wt}_M(s,t) \le 0 < 2 = y_s + y_t$.
- The vertex $u \neq d(s)$. 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 the definition of $\vec{y}$. Hence $\mathsf{wt}_M(s,t) = -2 < 0 = y_s + y_t$.

So in both cases, the edge (s,t) is slack with respect to M and its witness $\vec{y}$. 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. Consider again the two stable matchings M'_r and M'_ℓ defined earlier (see case 2) in $H^* = (A'_H \cup B'_H, E'_H)$. The vertex s is matched to t_r in M'_r and it is matched to its worst neighbor $d(s)$ in M'_ℓ.

In the graph H^* we have $t_r \succ_s t \succ_s d(s)$. We know that (s, t_r) and $(s, d(s))$ are stable edges in H^* since $(s, t_r) \in M'_r$ and $(s, d(s)) \in M'_\ell$. However, (s,t) is not a stable edge in H^* since (s,t) is not a popular edge 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 2).

The stable matching M' in H^* projects to a popular matching M in H; let $\vec{y} \in \{\pm 1\}^{2n}$ be M's witness as described earlier. There are again two subcases.

- The partner of t in M' is a *primed* vertex.[4] We have $y_s = y_t = 1$ by the definition of $\vec{y}$. Note that $\mathsf{wt}_M(s,t) \le 0$ since s prefers its partner in M to t. Hence $\mathsf{wt}_M(s,t) \le 0 < 2 = y_s + y_t$.
- The partner of t in M' is an *unprimed* vertex. We have $y_s = 1$ and $y_t = -1$ by the definition of $\vec{y}$. Both s and t prefer their respective partners in M to each other. Thus $\mathsf{wt}_M(s,t) = -2$. Hence $\mathsf{wt}_M(s,t) = -2 < 0 = y_s + y_t$.

[4] Recall that vertices in B'_H prefer any primed neighbor to any unprimed neighbor.

So in both cases, the edge (s,t) is slack with respect to M and its witness $\vec{y}$. 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. This finishes the proof of the lemma. □

3 The Fairly Popular Matching Polytope

We will prove Theorem 1 in this section. The high-level intuition for this proof is similar to that of Lemma 1. We would like to construct a new marriage instance G' (analogous to H^*) so that there is a surjective mapping from the set of stable matchings in G' to the set of fairly popular matchings in G. The key to this mapping in Section 2.2 was Theorem 4 (in fact, a sharper version from [22]).

Theorem 2 tells us that a matching M is fairly popular if and only if $\tilde{M}$, which is M augmented with self-loops at unmatched vertices, is a perfect matching in the graph G_p whose edge set is the set of popular fractional edges along with self-loops at unstable vertices. Thus, as done in Section 2.2, we can capture a fairly popular matching M as an optimal solution to a certain LP (see (LP3)). An optimal solution to the dual LP will be a dual certificate for M. We have a result analogous to Theorem 4 for fairly popular matchings (see Lemma 2).

As we will see, dual certificates for fairly popular matchings are more complicated than dual certificates for popular matchings. So rather than one marriage instance G', for each connected component C in G_p, we construct *two* instances G'_c and G''_c such that the restriction of any fairly popular matching M to the edge set of component C (call this matching M_c) can be realized as a stable matching either in instance G'_c or in instance G''_c. Thus we can compute a min-cost fairly popular matching as $M = \cup_C M_c$ for appropriate matchings M_c.

We will see the LP framework for fairly popular matchings in Section 3.1. Our characterization of fairly popular matchings is in Section 3.2. This characterization will be used in Section 3.3 to solve the min-cost fairly popular matching problem in polynomial time. Section 3.4 has the missing proofs from Section 3.3.

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 [10]) 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 2 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, we will use the edge weight function wt_M defined in Section 2.2. This function is now defined on $E \cup \{(u,u) : u \in A \cup B\}$ and we focus on the subset $\tilde{E}_p$. For any $(a,b) \in E_p$, we have $\mathsf{wt}_M(a,b) \in \{\pm 2, 0\}$ and for any unstable vertex u, we have $\mathsf{wt}_M(u,u) \in \{-1, 0\}$.

Consider the following linear program (LP3) analogous to (LP1) from Section 2.2. For each vertex v, let $\delta_p(v)$ be the set of edges incident to v in G_p.

$$\text{maximize} \sum_{e \in \tilde{E}_p} \mathsf{wt}_M(e) \cdot x_e \tag{LP3}$$

subject to

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

The above linear program computes a max-weight (wrt wt_M) perfect matching $\tilde{S}$ in G_p. It follows from Theorem 2 that S is a supporting matching. We have $\mathsf{wt}_M(\tilde{S}) = \Delta(S, M)$. Thus if the optimal value of (LP3) is positive then there exists a supporting matching that defeats M; else $\Delta(S, M) \leq 0$ for all supporting matchings S, so M is fairly popular.

For any stable matching S in G, note that $\Delta(S, M) \geq 0$. Since $\tilde{S} \subseteq \tilde{E}_p$, the optimal value of (LP3) has to be at least 0. Hence M is fairly popular if and only if the optimal value of (LP3) is 0.

Let $U \subseteq A \cup B$ be the set of unstable vertices in G. The linear program (LP4) is the dual LP.

$$\text{minimize} \sum_{v\in A\cup B} \alpha_v \tag{LP4}$$

subject to

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

Hence M is fairly popular if and only if there exists a feasible solution $\vec{\alpha}$ to (LP4) such that $\sum_{v\in A\cup B} \alpha_v = 0$.

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 $A_c = A \cap C$ is the same as the number of stable vertices in $B_c = B \cap C$. Hence there are k stable vertices in A_c if and only if there are k stable vertices in B_c.

Lemma 2. *A matching M is fairly popular if and only if there exists a feasible solution $\vec{\alpha}$ to* (LP4) *such that for every connected component C in G_p, we have $\sum_{v\in C} \alpha_v = 0$ and furthermore,*

- *either $\alpha_v \in \{0, \pm 2, \pm 4, \ldots, \pm(2k-2)\}$ 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.

We will first prove the following claim which will be used in the proof of Lemma 2. Let M be a fairly popular matching in G and let $\vec{\alpha}$ be an optimal solution to (LP4). The constraint matrix of (LP4) is totally unimodular, so we can assume that $\vec{\alpha} \in \mathbb{Z}^n$.

Claim 2 *For any connected component C in G_p, $\sum_{v\in C} \alpha_v = 0$. Furthermore, the α-values of all the vertices in C have the same parity.*

Proof. Let S be any stable matching in G. Let $S_c = S \cap (C \times C)$ and let $M_c = M \cap (C \times C)$. Since S_c is a stable matching in C, it is a popular matching in C; hence $\phi(S_c, M_c) \geq \phi(M_c, S_c)$. That is, $\Delta(S_c, M_c) \geq 0$ or equivalently, $\mathsf{wt}_{M_c}(\tilde{S}_c) = \mathsf{wt}_M(\tilde{S}_c) \geq 0$. Thus $\sum_{v\in C} \alpha_v \geq 0$.

Consider $\sum_C \sum_{v\in C} \alpha_v$ where the sum is over all connected components C in G_p. This sum equals $\sum_{v\in A\cup B} \alpha_v$. Since M is fairly popular, $\sum_{v\in A\cup B} \alpha_v = 0$. Since $\sum_{v\in C} \alpha_v \geq 0$ for each connected component C, it has to be the case that $\sum_{v\in C} \alpha_v = 0$ for each connected component C in G_p.

Every edge in E_p belongs to some popular fractional matching in G. Let $\vec{q}$ be the popular fractional matching that edge $(a, b) \in E_p$ belongs to, where a and b are in C. We have $\Delta(\vec{q}, M) = 0$ since $\vec{q}$ is a popular fractional matching, thus $\vec{q}$ is an optimal solution to (LP3). Because $\vec{\alpha}$ is an optimal solution to (LP4), 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. ◻

Proof of Lemma 2. Let M be a matching such that there exists a feasible solution $\vec{\alpha}$ to (LP4) 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 M is fairly popular.

Conversely, let M be a fairly popular matching in G and let $\vec{\alpha}$ be an integral optimal solution to (LP4). By Claim 2, $\sum_{v\in C} \alpha_v = 0$ and the α-values of all the vertices in C have the same parity.

Case 1: Suppose every vertex in C is stable. Then we can update the α-values of vertices in C as follows for any value t: let $\alpha_a = \alpha_a - t$ for all $a \in A_c$ and $\alpha_b = \alpha_b + t$ for all $b \in B_c$. The updated α-values are also a feasible solution to (LP4) since $\alpha_a + \alpha_b$ for any $(a, b) \in E_p$ (where $a, b \in C$) 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 $u \in C$.

The sum of α-values of all vertices in C is unchanged by this update since $|A_c| = |B_c| = k$ (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 by complementary slackness (see the proof of Claim 2), 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 A_c (and also in B_c) is k, we can conclude that there exists an optimal solution $\vec{\alpha}$ to (LP4) such that $\alpha_v \in \{0, \pm 2, \ldots, \pm(2k - 2)\}$ for all $v \in C$.

Case 2: Let us now assume that C has at least one unstable vertex. Consider the matching $\tilde{S} = S \cup \{(u, u) : u \in U\}$, where S is any stable matching in G and U is the set of unstable vertices in G. The matching $\tilde{S}$ is an optimal solution to (LP3). 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 2.1. 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 $\alpha_v \in \{0, \pm 2, \ldots, \pm(2k - 2)\}$ for all $v \in C$.

Case 2.2: 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 2, 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 $\vec{\gamma}$ (this is the vector $\vec{\alpha}$ in Lemma 2 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 $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ 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 $\vec{\gamma}$ such that $M_c = M \cap (C \times C)$ and $\vec{\gamma}$ satisfy properties 1-3 given above. Such a vector $\vec{\gamma}$ will be called a *witness* of M_c. Let $G_c = (C, E_c)$ where $E_c = E_p \cap (C \times C)$.

Definition 3. *Call a matching M_c in G_c* valid *if it has a witness, i.e., there exists a vector $\vec{\gamma}$ such that M_c and $\vec{\gamma}$ 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 2, $\mathcal{F}_c$ 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 $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ 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 $\vec{\gamma}$ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k-1)\}$ for all $v \in C$.

3.3 The fairly popular matching polytope

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 K be the set of stable vertices in G and let $|K \cap C| = 2k$.

A colorful multigraph. We will construct a multigraph G'_c on vertex set $A_c \cup B_c$. Its edge set E'_c is described below. Furthermore, each edge in E'_c has a *color* associated with it. Corresponding to every edge $(a, b) \in E_c$, the following parallel colored edges are in E'_c:

- If both a and b are in K (i.e., both are stable vertices in G) then there are $2k-1$ parallel edges (a, b) in E'_c. Each copy of the edge (a, b) has a distinct color in $\{0, \pm 1, \ldots, \pm(k-1)\}$ (see Fig. 2).

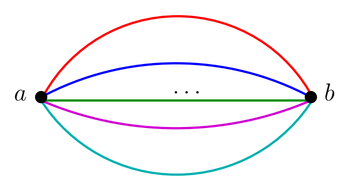

Fig. 2. G'_c has $2k-1$ parallel colored copies of $(a, b) \in E_c$ where a and b are stable vertices.

- If one of a, b is an unstable vertex in G then there is only one edge (a, b) in E'_c and it has color 0.

Since unstable vertices form an independent set, for any edge $(a, b) \in E_c$, note that at least one of a, b has to be in the set K of stable vertices. In the multigraph G'_c, the preference order of a vertex over its incident edges is as follows.

- each vertex in A prefers any lower colored edge to any higher colored edge;
- each vertex in B prefers any higher colored edge to any lower colored edge.

For any color i, the preference order of any vertex v among color i edges is exactly as per its preference order of the corresponding neighbors in G.

Stable matchings in G'_c. A matching N in the multigraph G'_c is a subset of E'_c such that each vertex in $A_c \cup B_c$ has at most one edge of N is incident to it. An edge $e = (a, b)$ (say, of color i) in G'_c *blocks* matching N if the following two conditions hold:

- Condition 1: (i) a is unmatched in N or (ii) a is matched in N along a color j edge where $j > i$ or (iii) a is matched in N along a color i edge to a neighbor worse than b.
- Condition 2: (i) b is unmatched in N or (ii) b is matched in N along a color j edge where $j < i$ or (iii) b is matched in N along a color i edge to a neighbor worse than a.

Matching N is *stable* in G'_c if there is no edge in E'_c that blocks N.

Valid matchings. Recall *valid* matchings in the instance G_c (see Definition 3). For any valid matching M_c in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$, define the matching M'_c in G'_c as follows.

- For every edge $(a, b) \in M_c$: include the edge (a, b) *colored* i in M'_c where $\gamma_b = 2i$.

We will show in Theorem 5 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 *colorless* M'_c, i.e., the colors of edges in M'_c are ignored. So M_c is a matching in G_c. Theorem 5 shows that M_c is a *valid* matching in G_c. The proof of Theorem 5 uses ideas from [26,29] and is given in Section 3.4.

Theorem 5. *M_c is a valid matching in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$ if and only if M'_c is a stable matching in G'_c.*

An extension of $\mathcal{F}^0_c$. For any vertex v in G'_c, let $\delta'_c(v)$ be the set of edges incident to v in G'_c. For any edge $(a, b) \in E_c$ and $i \in \{0, \pm 1, \ldots, \pm(k-1)\}$ such that there is an edge (a, b) colored i in G'_c, let $(a, b)_i$ denote the copy of the edge (a, b) colored i in G'_c.

For $v \in \{a, b\}$, let $\{e : e \succ_v (a, b)_i\} \subseteq \delta'_c(v)$ be the set of all edges in E'_c that v prefers to $(a, b)_i$. Consider constraints (1)-(2) in variables x_e where $e \in E'_c$ and λ_c (this variable will be defined later).

$$\sum_{e:\, e \succ_a (a,b)_i} x_e + \sum_{e':\, e' \succ_b (a,b)_i} x_{e'} + x_{(a,b)_i} \geq \lambda_c \quad \forall (a, b)_i \in E'_c \tag{1}$$

$$x_e \geq 0 \quad \forall e \in E'_c \qquad \text{and} \qquad \sum_{e \in \delta'_c(v)} x_e \leq \lambda_c \quad \forall v \in A'_c \cup B'_c. \tag{2}$$

The constraints in (2) with 1 replacing λ_c describe the matching polytope of G'_c and for each edge $(a, b)_i \in E'_c$, we get the stability constraint for edge $(a, b)_i$ by replacing λ_c with 1 in constraint (1). Thus constraints (1)-(2) with 1 replacing λ_c (wherever λ_c occurs) describe the stable matching polytope $\mathcal{S}'_c$ of G'_c (by [33]). There are several proofs of this and these proofs also hold for multigraphs—recall that G'_c is a multigraph. More concretely, it is easy to check that the simple proof given in [35, Theorem 1] holds for a multigraph.

By Theorem 5, the constraints formulating $\mathcal{S}'_c$ along with $x_{(a,b)} = \sum_i x_{(a,b)_i}$ for each edge (a, b) in E_c where i ranges over the colors of all the copies[5] of edge (a, b) in G'_c describe an extension of $\mathcal{F}^0_c$, where $\mathcal{F}^0_c$ is the convex hull of the edge incidence vectors of valid matchings in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$.

[5] So $i \in \{0, \pm 1, \ldots, \pm(k-1)\}$ if both a and b are stable vertices and $i = 0$ if one of them is unstable.

Another colorful multigraph. We will now construct another multigraph G''_c on vertex set $A_c \cup B_c$. Its edge set E''_c is defined below. As before, each edge in E''_c has a color associated with it. Corresponding to each edge $(a, b) \in E_c$, the following colored edges are in E''_c:

- If both a and b are in K then there are $2k$ parallel edges (a, b) in E'_c. Each copy of the edge (a, b) has a distinct color in $\{0, \pm 1, \ldots, \pm(k-1), k\}$.
- If one of a, b is an unstable vertex in G then there are *two* edges (a, b) in E'_c. One of these edges has color 0 and the other has color 1.

Regarding the preferences of a vertex over its incident edges, as before:

- each vertex in A prefers any lower colored edge to any higher colored edge;
- each vertex in B prefers any higher colored edge to any lower colored edge.

For any color i, the preference order of any vertex v among color i edges is exactly as per its preference order of the corresponding neighbors in G. The definition of a stable matching in the multigraph G''_c is the same as given earlier for multigraph G'_c.

For any valid matching M_c in G_c with a witness $\vec{\gamma}$ where $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k-1)\}$ for all $v \in C$, define the matching M''_c in G''_c as follows.

- For every edge $(a, b) \in M_c$: include the edge (a, b) *colored* i in M''_c where $\gamma_b = 2i - 1$.

We will show in Theorem 6 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 *colorless* M''_c, i.e., the colors of edges in M''_c are ignored. So M_c is a matching in G_c. Theorem 6 (proof given in Section 3.4) shows that M_c is a valid matching in G_c.

Theorem 6. *M_c is a valid matching in G_c with a witness $\vec{\gamma}$ 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.*

An extension of $\mathcal{F}^1_c$. For any vertex v in G''_c, let $\delta''_c(v)$ be the set of edges incident to v in G''_c. Consider constraints (3)-(4) in variables y_e where $e \in E''_c$ and λ_c.

$$\sum_{e:\, e \succ_a (a,b)_i} y_e + \sum_{e':\, e' \succ_b (a,b)_i} y_{e'} + \quad y_{(a,b)_i} \quad \geq \quad 1 - \lambda_c \quad \forall (a,b)_i \in E''_c \tag{3}$$

$$y_e \geq 0 \quad \forall e \in E''_c \qquad \text{and} \qquad \sum_{e \in \delta''_c(v)} y_e \;\leq\; 1 - \lambda_c \quad \forall v \in A''_c \cup B''_c. \tag{4}$$

Constraints (3)-(4) with 0 replacing λ_c (wherever λ_c occurs) describe the stable matching polytope $\mathcal{S}''_c$ of G''_c (by [33]). The stability constraint for edge $(a, b)_i$ in E''_c is given by (3) with 0 replacing λ_c and the constraints in (4) with 0 replacing λ_c describe the matching polytope of G''_c. By Theorem 6, the constraints formulating $\mathcal{S}''_c$ along with $y_{(a,b)} = \sum_i y_{(a,b)_i}$ for $(a, b) \in E_c$ where i ranges over the colors of all the copies[6] of edge (a, b) in G''_c describe an extension of $\mathcal{F}^1_c$. Recall that $\mathcal{F}^1_c$ is the convex hull of the edge incidence vectors of valid matchings in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k-1)\}$ for all $v \in C$.

[6] So $i \in \{0, \pm 1, \ldots, \pm(k-1), k\}$ if both a and b are stable vertices and $i \in \{0, 1\}$ if one of them is unstable.

The valid matching polytope. We know from Lemma 2 that any valid matching in C has a witness $\vec{\gamma}$ where either (i) $\gamma_v \in \{0, \ldots, \pm(2k-2)\}$ 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 $\mathcal{F}_c^0 \cup \mathcal{F}_c^1$ is the valid matching polytope $\mathcal{F}_c$ of G_c.

Balas' theorem [2] on the convex hull of $\mathcal{F}_c^0 \cup \mathcal{F}_c^1$ says this polytope is described as follows:

$$\mathsf{conv}(\mathcal{F}_c^0 \cup \mathcal{F}_c^1) = \{z : \exists (x, y, \lambda_c) \text{ such that } z = x\lambda_c + y(1-\lambda_c) \text{ where } x \in \mathcal{F}_c^0, y \in \mathcal{F}_c^1, \text{ and } 0 \le \lambda_c \le 1\}.$$

Thus the variable $\lambda_c \in [0, 1]$ gets introduced and we get constraints (1)-(6) where constraints (1)-(4) are given above and constraints (5)-(6) are given below.

$$z_{(a,b)} = x_{(a,b)} + y_{(a,b)} \quad \forall (a, b) \in E_c \tag{5}$$

$$z_e = 0 \quad \forall e \in (E \cap (C \times C)) \setminus E_c \quad \text{and} \quad 0 \le \lambda_c \le 1 \tag{6}$$

Hence the polytope defined by (1)-(6) is an extension of the polytope $\mathcal{F}_c$. Thus Theorem 7 stated below follows.

Theorem 7. *The polytope $\mathcal{P}_c$ defined by constraints (1)-(6) 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 $z_e = 0$ 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 1 stated in Section 1.

3.4 Proofs of Theorem 5 and Theorem 6

We will first prove Theorem 5. This will be proved in two parts: Lemma 3 and Lemma 4.

Lemma 3. *Let M_c be a valid matching in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$. Then M'_c is a stable matching in G'_c.*

Before we prove the above lemma, we will prove the following claim. Recall that K (resp., U) is the set of stable (resp., unstable) vertices in G.

Claim 3 *All vertices in $K \cap C$ are matched in M_c and no vertex in $U \cap C$ is matched in M_c.*

Proof. Consider (LP3) 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 at most 0 since there exists a dual feasible solution $\vec{\gamma}$ with $\sum_{u \in C} \gamma_u = 0$ (recall that $\vec{\gamma}$ obeys properties 1-3). This means no supporting matching in G_c defeats M_c, so M_c is fairly popular in G_c and thus it is a supporting matching. So M_c has to match all stable vertices in G_c (by Theorem 2). The set of stable vertices in G_c is $K \cap C$ since the set of vertices matched in the stable matching S_c in G_c is $K \cap C$, where S is any stable matching in G and $S_c = S \cap (C \times C)$.

We now need to show that no vertex in $U \cap C$ is matched in M_c. Observe that $\mathsf{wt}_M(\tilde{S}_c) = \Delta(S_c, M_c) = 0$ (since S_c is popular in G_c). So $\tilde{S}_c$ is an optimal solution to (LP3) with M_c replacing M and $\tilde{E}_c$ replacing $\tilde{E}$. For any $u \in U \cap C$, the self-loop $(u, u) \in \tilde{S}_c$. Since $\vec{\gamma}$ is an optimal solution to the dual LP, 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 has to be the case that $\gamma_u = \mathsf{wt}_{M_c}(u, u) = 0$, i.e., u is left unmatched in M_c. ◻

We are now ready to prove Lemma 3. We need to show the matching M'_c is stable in the colorful graph G'_c. Hence for any edge (a, b) in G_c and any color i such that $(a, b)_i$ is present in G'_c,[7] we need to show the edge $(a, b)_i$ does not block M'_c. There are three cases based on the values of γ_a and γ_b: in each case we show none of the parallel edges $(a, b)_i$ blocks M'_c.

Proof (of Lemma 3). Observe that $\tilde{M}_c$ is an optimal solution to (LP3) with M_c replacing M and $\tilde{E}_c$ replacing $\tilde{E}$. Hence for any $(s, t) \in M_c$, we have $\gamma_s + \gamma_t = \mathsf{wt}_{M_c}(s, t) = 0$ by complementary slackness. Consider any edge (a, b) in G_c.

Suppose $a \in U \cap C_A$. Then $\gamma_a = 0$ (see the proof of Claim 3). For any $(a, b) \in E_c$, we have $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a, b) \geq 0$. So $\gamma_b \geq 0$. If $\gamma_b = 0$ then $\mathsf{wt}_{M_c}(a, b) = 0$. This means $(z, b)_0 \in M'_c$ for some neighbor z that b prefers to a. Else $\gamma_b > 0$ and so $(z, b)_i \in M'_c$ for some edge (z, b) incident to b. We have $2i = \gamma_b > 0$, thus $i > 0$. Recall that b prefers any positive color edge to a color 0 edge. Thus b is matched along an edge $(z, b)_i$ that it prefers to $(a, b)_0$, hence $(a, b)_0$ does not block M'_c.

Suppose $a \in K \cap C_A$. We will show no edge $(a, b)_\ell$ in G_c blocks M'_c where $\ell \in \{0, \pm 1, \ldots, \pm(k-1)\}$ for stable b and $\ell = 0$ for unstable b. Since a is a stable vertex, we know that a is matched in M_c (by Claim 3). Let $\gamma_a = -2i$. So $(a, w)_i \in M_c$ for some neighbor w of a. Let $\gamma_b = 2j$. We know that $\gamma_a + \gamma_b = -2i + 2j \geq \mathsf{wt}_M(a, b)$. Since $\mathsf{wt}_M(a, b) \geq -2$, it follows that $j \geq i - 1$. Let us consider the following three cases.

1. $j = i - 1$: This means that $\gamma_a + \gamma_b = -2i + 2(i-1) = -2 \geq \mathsf{wt}_{M_c}(a, b)$. So $\mathsf{wt}_{M_c}(a, b) = -2$, i.e., both a and b prefer their partners in M_c to each other. Thus b is matched in M_c to a neighbor z that it prefers to a. So the edge $(z, b)_{i-1} \in M'_c$, where b prefers $(z, b)_{i-1}$ to $(a, b)_{i-1}$. Moreover, a prefers w to b. Hence the edge $(a, w)_i \in M'_c$, where a prefers $(a, w)_i$ to $(a, b)_i$. Thus neither $(a, b)_{i-1}$ nor $(a, b)_i$ blocks M'_c.
   Furthermore, a prefers any lower color edge to any higher color edge—so a prefers $(a, w)_i$ to $(a, b)_\ell$ for all $\ell > i$. Similarly, b prefers any higher color edge to any lower color edge—so b prefers $(z, b)_{i-1}$ to $(a, b)_\ell$ for all $\ell < i - 1$. Hence no edge $(a, b)_\ell$ in G_c blocks M'_c.
2. $j = i$: This means that $\gamma_a + \gamma_b = -2i + 2i = 0 \geq \mathsf{wt}_{M_c}(a, b)$. So $\mathsf{wt}_{M_c}(a, b) \leq 0$. Thus either $(a, b)_i \in M'_c$ or one of a, b prefers the edge along which it is matched in M_c to $(a, b)_i$. So the edge $(a, b)_i$ does not block M'_c in either case.
   Suppose b is a stable vertex (recall that a is a stable vertex). Then (a, w) and (z, b) are in M_c, where $w = b$ and $z = a$ if $(a, b) \in M_c$. Thus $(a, w)_i$ and $(z, b)_i$ are in M'_c. Since a prefers any lower color edge to any higher color edge, a prefers $(a, w)_i$ to $(a, b)_\ell$ for all $\ell > i$. Similarly, b prefers $(z, b)_i$ to $(a, b)_\ell$ for all $\ell < i$. Hence no edge $(a, b)_\ell$ in G'_c blocks M'_c.
   If b is unstable then $i = 0$ and $(a, b)_0$ is the only edge in G'_c between a and b. Moreover, since $\mathsf{wt}_{M_c}(a, b) \leq 0$, the vertex a prefers its partner w to b, hence a prefers the edge $(a, w)_0$ to $(a, b)_0$. Thus $(a, b)_0$ does not block M'_c.
3. $j \geq i + 1$: If b is an unstable vertex then $(a, w)_i \in M'_c$ where $i \leq -1$ (since $j = 0$). Since a prefers lower color edges to higher color edges, a prefers $(a, w)_i$ to $(a, b)_0$.
   Suppose b is a stable vertex. Because a is a stable vertex, we have (a, w) and (z, b) in M_c; so $(a, w)_i$ and $(z, b)_j$ are in M'_c. Since a prefers lower color edges to higher color edges, a prefers $(a, w)_i$ to $(a, b)_\ell$ for $\ell \geq i + 1$. Similarly, b prefers $(z, b)_j$ to $(a, b)_\ell$ for $\ell \leq j - 1$. Thus no edge $(a, b)_\ell$ where $\ell \in \{0, \pm 1, \ldots, \pm(k-1)\}$ blocks M'_c.

Thus we have shown that M'_c is a stable matching in G'_c. □

[7] Recall that $i \in \{0, \pm 1, \ldots, \pm(k-1)\}$ if both a and b are stable vertices, else $i = 0$.

We now prove the converse of Lemma 3, i.e., we show the *colorless* matching M_c obtained from the stable matching M'_c in G'_c is a valid matching in G_c. This involves defining a witness $\vec{\gamma}$ for M_c. We will use the color of the edge along which a vertex is matched in M'_c to define its γ-value. The non-trivial step is to show that to show every $(a, b) \in E_c$ is *covered*, i.e., $\gamma_a + \gamma_b \geq \mathsf{wt}_{M_c}(a, b)$.

Lemma 4. *If M'_c is a stable matching in G'_c then M_c (i.e., the colorless M'_c) is a valid matching in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$.*

Proof. Let S be any stable matching in G and let $S_c = S \cap (C \times C)$. It is easy to check that $S'_c = \{(a, b)_0 : (a, b) \in S_c\}$ is a stable matching in G'_c. The set of vertices left unmatched in S'_c is $\{u : u \in U \cap C\}$. All stable matchings in G'_c match the same subset of vertices—this fact is well-known for simple graphs [18] and it holds for multigraphs as well (recall that G'_c is a multigraph). For the sake of completeness, we include a proof of this fact for multigraphs as Proposition 3 in the appendix.

Since M'_c is a stable matching in G'_c, it matches all vertices of G'_c except the vertices u where $u \in U \cap C$. In order to prove that M_c is a valid matching in G_c, we define $\vec{\gamma}$ as follows:

- for every vertex $u \in U \cap C$: let $\gamma_u = 0$;
- for every edge $(s, t)_i \in M'_c$ where $s \in A_c$ and $t \in B_c$: let $\gamma_s = -2i$ and $\gamma_t = 2i$.

Since $i \in \{0, \pm 1, \ldots, \pm(k-1)\}$, it follows that $\gamma_v \in \{0, \pm 2, \ldots, \pm(2k-2)\}$ for all $v \in C$. For any vertex $u \in U \cap C$ (each such vertex is unmatched in M_c), we have $\gamma_u = 0 = \mathsf{wt}_{M_c}(u, u)$. We also have $\sum_{v \in C} \gamma_v = \sum_{(s,t) \in M_c} (\gamma_s + \gamma_t) = 0$.

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 in the definition of witness hold and thus M_c is a valid matching in G_c with $\vec{\gamma}$ as a witness. Suppose $\gamma_a = -2i$ and $\gamma_b = 2j$. We need to show that $-2i + 2j \geq \mathsf{wt}_{M_c}(a, b)$. Consider the following four 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 $e \in E_c$.
2. $j = i$: Since the edge $(a, b)_i$ does not block M'_c, either $(a, b)_i \in M'_c$ or one of a, b is matched along an edge that it prefers to $(a, b)_i$. Recall that the preference order of any vertex along color i edges is exactly as per its preference order of the corresponding neighbors 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. So $\mathsf{wt}_{M_c}(a, b) \leq 0$. Hence $\gamma_a + \gamma_b = -2i + 2i = 0 \geq \mathsf{wt}_{M_c}(a, b)$.
3. $j = i - 1$: If a is an unstable vertex then $(a, b)_0$ blocks M'_c since $(z, b)_{-1} \in M'_c$ for some neighbor z (recall that b prefers any color 0 edge to a color -1 edge). This contradicts the stability of M'_c, thus a is a stable vertex; so $(a, w)_i \in M'_c$ for some neighbor w. Also, $(z, b)_{i-1} \in M'_c$ for some neighbor z that b prefers to a. Otherwise the edge $(a, b)_{i-1}$ would block M'_c as a prefers any color $i - 1$ edge to any color i edge.
   Furthermore, b prefers $(a, b)_i$ to $(z, b)_{i-1}$ since b prefers any color i edge to any color $(i-1)$ edge. Since $(a, b)_i$ does not block M'_c, it has to be the case 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. Either b is unmatched (so b is unstable) or $(z, b)_j \in M'_c$. In the former case, the edge $(a, b)_0$ blocks M'_c since a is matched along a color $i \geq 2$ edge. In the latter case, the edge $(a, b)_{i-1}$ blocks M'_c. So in either case, M'_c has a blocking edge—a contradiction to its stability in G'_c. Thus we cannot have $j \leq i - 2$. □

Lemma 3 and Lemma 4 imply Theorem 5. We will now prove Theorem 6. This will again be proved in two parts: Lemma 5 and Lemma 6.

Lemma 5. *If M_c is a valid matching in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{\pm 1, \pm 3, \ldots, \pm(2k-1)\}$ for all $v \in C$ then M_c'' is a stable matching in G_c''.*

Before we prove the above lemma, we will prove the following claim.

Claim 4 *All vertices in C are matched in M_c.*

Proof. As shown in the proof of Claim 3, M_c is a supporting matching in G_c. Thus M_c has to match all stable vertices in G_c (by Theorem 2). For any unstable vertex u in C, the self-loop $(u,u) \in \tilde{S}_c$, where S is any stable matching in G and $S_c = S \cap (C \times C)$. Hence the constraint $\gamma_u \geq \mathsf{wt}_{M_c}(u,u)$ is tight (by complementary slackness). Because $\mathsf{wt}_{M_c}(u,u) \in \{0,-1\}$ and γ_u is odd, it has to be the case that $\gamma_u = \mathsf{wt}_{M_c}(u,u) = -1$, i.e., u is matched in M_c. □

The proof of Lemma 5 is similar to the proof of Lemma 3. In fact, this proof is simpler since there are no vertices left unmatched in M_c (by Claim 4).

Proof (of Lemma 5). We need to show that M_c'' is stable in the colorful graph G_c''. Hence for any edge (a,b) in G_c and any color ℓ such that $(a,b)_\ell$ is present in G_c',[8] we need to show that $(a,b)_\ell$ does not block M_c''. Since $\tilde{M}_c$ is an optimal solution to (LP3), for any $(s,t) \in M_c$, we have $\gamma_s + \gamma_t = \mathsf{wt}_{M_c}(s,t) = 0$ (by complementary slackness).

Let us now show that no edge $(a,b)_\ell$ in G_c blocks M_c''. Let $\gamma_a = -(2i-1)$. So $(a,w)_i \in M_c''$ for some neighbor w of a. Let $\gamma_b = 2j-1$. We know that $\gamma_a + \gamma_b = -2i+1+2j-1 \geq \mathsf{wt}_M(a,b) \geq -2$. Thus it follows that $j \geq i-1$. Let us consider the following three cases.

1. $j = i-1$: This means that $\gamma_a + \gamma_b = -2i+1+2(i-1)-1 = -2 \geq \mathsf{wt}_{M_c}(a,b)$. So $\mathsf{wt}_{M_c}(a,b) = -2$, i.e., both a and b prefer their partners in M_c to each other. Thus b has to be matched in M_c to a neighbor z that it prefers to a. So the edge $(z,b)_{i-1} \in M_c''$, where b prefers $(z,b)_{i-1}$ to $(a,b)_{i-1}$ in G_c''. Moreover, a prefers w to b. So the edge $(a,w)_i \in M_c''$, where a prefers $(a,w)_i$ to $(a,b)_i$. Hence neither $(a,b)_{i-1}$ nor $(a,b)_i$ blocks M_c''.
   Furthermore, a prefers any lower color edge to any higher color edge—so a prefers $(a,w)_i$ to $(a,b)_\ell$ for all $\ell > i$. Similarly, b prefers any higher color edge to any lower color edge—so b prefers $(z,b)_{i-1}$ to $(a,b)_\ell$ for all $\ell < i-1$. Hence no edge $(a,b)_\ell$ in G_c blocks M_c''.
2. $j = i$: This means that $\gamma_a + \gamma_b = -2i+1+2i-1 = 0 \geq \mathsf{wt}_{M_c}(a,b)$. So $\mathsf{wt}_{M_c}(a,b) \leq 0$. Thus (i) $(a,b)_i \in M_c''$ or (ii) a prefers w to b where $(a,w)_i \in M_c''$ or (iii) b prefers z to a where $(z,b)_i$ in M_c''. So the edge $(a,b)_i$ does not block M_c'' in any case.
   Since a prefers any lower color edge to any higher color edge, a prefers $(a,w)_i$ to $(a,b)_\ell$ for all $\ell > i$. Similarly, b prefers $(z,b)_i$ to $(a,b)_\ell$ for all $\ell < i$. Hence no edge $(a,b)_\ell$ in G_c'' blocks M_c''.
3. $j \geq i+1$: Suppose (a,w) and (z,b) are in M_c. So $(a,w)_i$ and $(z,b)_j$ are in M_c''. Since a prefers lower color edges to higher color edges, a prefers $(a,w)_i$ to $(a,b)_\ell$ for $\ell \geq i+1$. Similarly, b prefers $(z,b)_j$ to $(a,b)_\ell$ for $\ell \leq j-1$. Thus no edge $(a,b)_\ell$ blocks M_c''.

Thus we have shown that M_c'' is a stable matching in G_c''. □

[8] Recall that $\ell \in \{0, \pm 1, \ldots, \pm(k-1), k\}$ if both a and b are stable vertices, else $\ell \in \{0,1\}$.

Lemma 6. *If M_c'' is a stable matching in G_c'' then M_c is a valid matching in G_c with a witness $\vec{\gamma}$ such that $\gamma_v \in \{\pm 1, \pm 3, \dots, \pm(2k-1)\}$ for all $v \in C$.*

Before we prove the above lemma, we will prove the following claim.

Claim 5 *Any stable matching in G_c'' matches all vertices in C.*

Proof. Consider the subgraph G_c^0 of G_c'' with vertex set $C = A_c \cup B_c$ and edge set E_c^0 which is E_c'' restricted to color 0 and color 1 edges. So every adjacent pair of vertices in G_c^0 is connected by two parallel edges: one colored 0 and the other colored 1. As was the case in G_c'', every vertex in A_c prefers any color 0 edge to any color 1 edge while any vertex in B_c prefers any color 1 edge to color 0 edge. Among color i edges incident to any vertex v (where $i \in \{0, 1\}$), it is v's original preference order.

It follows from [24] that any stable matching in G_c^0 projects to a max-size popular matching in G_c, i.e., ignoring edge colors in any stable matching in G_c^0 yields a max-size popular matching (let P_c be such a matching) in G_c. Any vertex left unmatched in P_c has to be isolated in G_c (see Claim 6 in the appendix). Recall that C is a connected component of G_p and $|C| \geq 2$. Hence every vertex in C has at least one edge incident to it in G_p, and thus in G_c. Thus no vertex in C is left unmatched in the matching P_c.

Hence the original stable matching S_c^0 (whose colorless version is P_c) matches all vertices in G_c^0. We claim that any stable matching S_c^0 in G_c^0 is also a stable matching in G_c''. All color 0 and color 1 edges of G_c'' are in G_c^0 and each of the edges in $G_c'' \setminus G_c^0$ has either a color higher than 1 or a color lower than 0 and is between an adjacent pair in G_c^0.

Recall that any vertex in A_c prefers being matched along a color 0 or color 1 edge to being matched along a higher color edge while any vertex in B_c prefers being matched along a color 0 or color 1 edge to being matched along a lower color edge. Since all vertices in C are matched in S_c^0, none of the new edges in $G_c'' \setminus G_c^0$ blocks S_c^0. Thus the perfect matching S_c^0 is stable in G_c''. Because all stable matchings in G_c'' match the same subset of vertices (see Proposition 3 in the appendix), any stable matching in G_c'' matches all vertices in C. □

Proof (of Lemma 6). M_c'' is a stable matching in G_c''. By Claim 5, M_c'' matches all vertices in C. In order to prove that M_c is a valid matching in G_c, we will define $\vec{\gamma}$ as follows:

- for every edge $(s,t)_i \in M_c''$, let $\gamma_s = -(2i-1)$ and $\gamma_t = 2i - 1$.

Since $i \in \{0, \pm 1, \dots, \pm(k-1), k\}$, we have $\gamma_v \in \{\pm 1, \pm 3, \dots, \pm(2k-1)\}$ for all $v \in C$. We also have $\sum_{v \in C} \gamma_v = \sum_{(s,t) \in M_c} (\gamma_s + \gamma_t) = 0$. Furthermore, for any unstable vertex u, we have $(u,v)_i \in M_c''$ where v is a neighbor of u and $i \in \{0, 1\}$. Thus $|\gamma_u| = |2i - 1|$ where $i \in \{0, 1\}$. So $\gamma_u \in \{\pm 1\}$, in other words, $\gamma_u \geq -1 = \mathsf{wt}_{M_c}(u,u)$.

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 in the definition of witness hold and thus M_c is valid in G_c with $\vec{\gamma}$ as a witness. Suppose $\gamma_a = -(2i-1)$ and $\gamma_b = 2j - 1$. Let us consider the following four cases:

1. $j \geq i + 1$: So $\gamma_a + \gamma_b = -2i + 1 + 2j - 1 = 2(j - i) \geq 2$. Since $\mathsf{wt}_{M_c}(e) \in \{\pm 2, 0\}$ for any $e \in E$, we have $\mathsf{wt}_{M_c}(a,b) \leq 2 \leq \gamma_a + \gamma_b$.
2. $j = i$: Since the edge $(a,b)_i$ does not block M_c'', either $(a,b)_i \in M_c''$ or one of a, b is matched along an edge preferred to $(a,b)_i$. 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 $\mathsf{wt}_{M_c}(a,b) \leq 0$. Hence $\gamma_a + \gamma_b = -(2i-1) + 2i - 1 = 0 \geq \mathsf{wt}_{M_c}(a,b)$.

3. $j = i - 1$: So $(a, w)_i$ and $(z, b)_{i-1}$ are in M_c''. Recall that b prefers any higher color edge to any lower color edge, thus b prefers $(a, b)_i$ to $(z, b)_{i-1}$. Since the edge $(a, b)_i$ does not block M_c'' (because M_c'' is a stable matching in G_c''), it has to be the case that a prefers $(a, w)_i$ to $(a, b)_i$, in other words, a prefers w to b.

   Similarly, a prefers any lower color edge to any higher color edge, thus a prefers $(a, b)_{i-1}$ to $(a, w)_i$. Since the edge $(a, b)_{i-1}$ does not block M_c'', it has to be the case that b prefers $(z, b)_{i-1}$ to $(a, b)_{i-1}$, in other words, b prefers z to a. 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. $j \le i - 2$: We have $(a, w)_i$ and $(z, b)_j$ in M_c'' where $j \le i - 2$. So the edge $(a, b)_{i-1}$ blocks M_c''. This contradicts the stability of M_c'' in G_c''. Thus this case does not occur. □

This finishes the proof of Theorem 6.

4 A Hardness Result

We prove Proposition 1 and Theorem 3 in this section. Let $\tilde{G} = (A \cup B, \tilde{E})$ where $\tilde{E} = E \cup \{(u, u) : u \in A \cup B\}$. Thus we can regard any fractional matching $\vec{x}$ in G as a perfect fractional matching in $\tilde{G}$ by setting $x_{(u,u)} = 1 - \sum_{e \in \delta(u)} x_e$ for all vertices u.

Let $\mathcal{M}_G$ be the matching polytope of the bipartite graph $G = (A \cup B, E)$. Any popular matching M satisfies $\Delta(\vec{x}, M) \le 0$ for all $\vec{x} \in \mathcal{M}_G$ where $\Delta(\vec{x}, M) = \mathsf{wt}_M(\vec{x}) = \sum_{e \in \tilde{E}} \mathsf{wt}_M(e) \cdot x_e$. Note that the constraint $\Delta(\vec{x}, M) \le 0$ involves $m + n$ variables x_e for $e \in \tilde{E}$, where $|A \cup B| = n$ and $|E| = m$. By substituting $x_{(u,u)} = 1 - \sum_{e \in \delta(u)} x_e$ for every vertex u, this constraint involves only the m variables x_e for $e \in E$.

Observation 1 *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 $\vec{x} \in \mathcal{M}_G$ satisfies $\Delta(\vec{x}, M) \le 0$ for all popular matchings M. So the intersection of $\mathcal{M}_G$ with the constraints $\Delta(\vec{x}, M) = 0$ for all popular matchings M 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 $\mathcal{X}$:

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

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 $\Delta(\vec{x}, M) = 0$ are necessary and the rest are redundant. Thus there exist at most $t \le m$ popular matchings $M_1, \ldots, M_t$ such that if a matching N satisfies $\Delta(N, M_i) = 0$ for $1 \le i \le t$ then the edge incidence vector of N belongs to $\mathcal{X}$, i.e., such a matching N is not defeated by any popular matching. Hence Proposition 1 follows.

The NP-hardness proof. We now prove Theorem 3 which states that in spite of the compactness result given by Proposition 1, it is NP-complete to decide if there exists a popular matching that defeats a given matching N. 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 a formula ψ, to decide if ψ is 1-in-3 satisfiable is NP-complete [34]. Given ψ, as done in [13], we will construct an instance G described below. The graph G has many gadgets. There is one gadget corresponding to each variable and several gadgets corresponding to each clause in G.

- The gadget for variable X_i is on 4 vertices x_i, y_i, x'_i, y'_i and is illustrated on the right in Fig. 3.
- Other than the clause and variable gadgets, there is one special gadget on four vertices a_0, b_0, z', z. The gadget formed by these four vertices is illustrated on the left in Fig. 3.

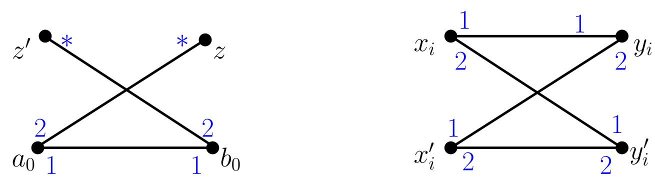

Fig. 3. The numbers on edges denote preferences: 1 is top choice, 2 is second choice, and $*$ denotes a number $\gg 1$. Vertices a_0, b_0, z', z form a single gadget on the left and the gadget corresponding to variable X_i is on the right.

There are several inter-gadget edges, i.e., edges with endpoints in different gadgets. However all inter-gadget edges are *unpopular* [13, Theorem 3.2]. Equivalently, the endpoints of every popular edge in G are within the same gadget. So any popular matching P in G contains either the pair $(a_0, z), (z', b_0)$ or the single edge (a_0, b_0).

Furthermore, there are two alternatives for the popular matching P from each variable gadget. Let n be the number of variables in ψ. For $i \in \{1, \ldots, n\}$:

- The popular matching P contains either the pair $(x_i, y_i), (x'_i, y'_i)$ or the pair $(x_i, y'_i), (x'_i, y_i)$.
  - If $\{(x_i, y_i), (x'_i, y'_i)\} \subseteq P$ then the gadget corresponding to X_i is in *zero* state in P.
  - If $\{(x_i, y'_i), (x'_i, y_i)\} \subseteq P$ then the gadget corresponding to X_i is in *unit* state in P.

The following theorem is Theorem 3.4 combined with Theorem 3.5 from [13].

Theorem 8 ([13]). *G has a popular matching P that matches all vertices except z and z' if and only if for each clause C in ψ, there is exactly one variable in C whose gadget is in unit state in P.*

The gadget for X_i being in unit state is interpreted as the variable X_i being set to *true* and this gadget being in zero state is interpreted as the variable X_i being set to *false*. Thus by Theorem 8, G has a popular matching that matches all vertices except z and z' if and only if ψ is 1-in-3 satisfiable. Hence it is NP-hard to decide if G has a popular matching that matches all vertices except z and z'.

The augmented instance G. In order to prove another hardness result on popular matchings, the reduction in [13] augments the above instance G with a gadget on four new vertices x_0, y_0, x'_0, y'_0. The edges within this new gadget are similar to those within any variable gadget (see the gadget on the right in Fig. 4). Inter-gadget edges are incident to the vertices x'_0, y'_0, however x_0, y_0 have no neighbors outside their gadget in the construction in [13]. As before, no inter-gadget edge belongs to any popular matching (see the proof of [13, Theorem 4.1]). So any popular matching in G contains either the pair $(x_0, y_0), (x'_0, y'_0)$ or the pair $(x_0, y'_0), (x'_0, y_0)$.

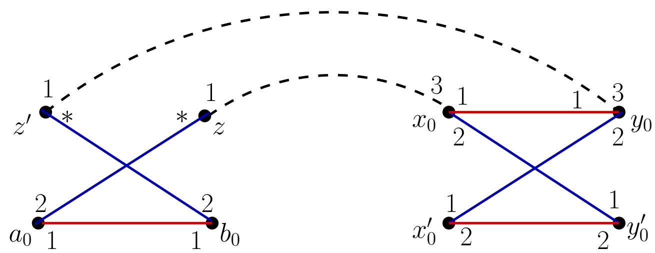

Fig. 4. The dashed edges do not belong to the instance used in [13] but these edges will be useful to us here.

It was shown in [13] that any popular matching in G that contains both (a_0, b_0) and (x_0, y'_0) has to match all vertices in G except z and z'. Thus by Theorem 8, it is NP-hard to decide if there exists a popular matching in G that contains both (a_0, b_0) and (x_0, y'_0). The following result is [13, Theorem 4.1(ii)].

Theorem 9 ([13]). *The instance G admits a popular matching that contains the two edges (a_0, b_0) and (x_0, y'_0) if and only if ψ is 1-in-3 satisfiable.*

We will use a slightly modified version of the above instance G to show the NP-hardness of deciding if there exists a popular matching that defeats a given matching N. We will add the dashed edges (x_0, z) and (z', y_0) to G (see Fig. 4). The top choices of z and z' are x_0 and y_0, respectively. The vertices x_0 and y_0 regard z and z' as their worst neighbors, respectively.

Observation 2 *Neither (x_0, z) nor (z', y_0) is a popular edge in G.*

Proof. Any popular matching in G that contains either (x_0, z) or (z', y_0) has to contain all the three edges $(x_0, z), (z', y_0), (x'_0, y'_0)$. Thus what we need to check is the following: there is *no* popular matching in G that contains the three edges $(x_0, z), (z', y_0), (x'_0, y'_0)$.

A matching N that contains these three edges is not popular since the matching M obtained from N by replacing these three edges with the two edges (x_0, y'_0) and (x'_0, y_0) is more popular. Observe that $\Delta(M, N) = 4 - 2 = 2$ since x_0, y_0, x'_0, y'_0 prefer M to N while z, z' (these are unmatched in M) prefer N to M and all other vertices are indifferent between M and N. □

We will now check that Theorem 9 continues to hold in this instance G augmented with the edges (x_0, z) and (z', y_0). Any popular matching M that contains (a_0, b_0) and (x_0, y'_0) has to leave z and z' unmatched. This is because $(a_0, z), (z', b_0)$ are the only popular edges incident to z, z' and since $(a_0, b_0) \in M$, neither of these edges belongs to M.

A dual witness $\vec{\alpha}$ (see Theorem 4) of such a popular matching M has to satisfy (i) $\alpha_{x_0} = \alpha_{y_0} = 1$ since $\alpha_{x_0} + \alpha_{y_0} \ge \mathsf{wt}_M(x_0, y_0) = 2$ because (x_0, y_0) is a blocking edge to M and (ii) $\alpha_z = \alpha_{z'} = 0$ since $\alpha_z = \mathsf{wt}_M(z, z) = 0$ and $\alpha_{z'} = \mathsf{wt}_M(z', z') = 0$ because z and z' are unmatched in M. As shown in [13], such a dual certificate $\vec{\alpha}$ will lead to a 1-in-3 satisfying assignment for ψ. The proof of [13, Theorem 4.1] uses $\vec{\alpha}$ to show that M has to match all vertices in G other than z and z' and the proof of [13, Theorem 3.4] uses M to define a truth assignment to the variables in ψ so that ψ is 1-in-3 satisfiable.

Conversely, if ψ is 1-in-3 satisfiable then as done in the proof of [13, Theorem 3.5], this satisfying assignment can be used to construct a popular matching M in the *old* instance G, i.e., without the edges $(x_0, z), (z', y_0)$, such that $\{(a_0, b_0), (x_0, y'_0)\} \subseteq M$. The dual certificate $\vec{\alpha}$ constructed in this proof satisfies $\alpha_{x_0} = \alpha_{y_0} = 1$ and $\alpha_z = \alpha_{z'} = 0$. Thus we have $\alpha_{x_0} + \alpha_z = 1 > 0 = \mathsf{wt}_M(x_0, z)$ and $\alpha_{z'} + \alpha_{y_0} = 1 > 0 = \mathsf{wt}_M(z', y_0)$. So these two edges are also covered by $\vec{\alpha}$. Hence M is popular in our new instance G (by Theorem 4).

Let $S = \{a_0, b_0, z', z, x_0, y_0, x'_0, y'_0\}$. Define N as follows:

$$N = N_0 \cup N_1 \text{ where } N_1 = \{(a_0, b_0), (x_0, z), (z', y_0), (x'_0, y'_0)\} \text{ and } N_0 \text{ is any stable matching in the subgraph of } G \text{ induced on} (A \cup B) \setminus S.$$

We know from Observation 2 that N is not popular. The non-trivial question is whether there is a popular matching more popular than N.

Lemma 7. *There exists a popular matching in G that defeats N if and only if ψ is 1-in-3 satisfiable.*

Proof. Let G_1 be the subgraph of G induced on $S = \{a_0, b_0, z, z', x_0, y_0, x'_0, y'_0\}$ and let G_0 be the subgraph induced on $(A \cup B) \setminus S$.

(The $\Rightarrow$ direction.) Suppose there is a popular matching M that is more popular than N. As mentioned earlier, no edge between G_0 and G_1 belongs to any popular matching. Hence $M = M_0 \cup M_1$ where M_i is within G_i, for $i = 0, 1$. Since M is popular in G, the matchings M_0 and M_1 have to be popular in G_0 and G_1, respectively. We have $\Delta(M, N) = \Delta(M_0, N_0) + \Delta(M_1, N_1)$. Moreover, $\Delta(M_0, N_0) = 0$ because M_0 and N_0 are popular matchings in G_0. Since $\Delta(M, N) > 0$, it must be the case that $\Delta(M_1, N_1) > 0$.

The graph G_1 has three popular matchings. These are $P_1 = \{(a_0, b_0), (x_0, y'_0), (x'_0, y_0)\}$, $P_2 = \{(a_0, b_0), (x_0, y_0), (x'_0, y'_0)\}$, and $P_3 = \{(a_0, z), (z', b_0), (x_0, y'_0), (x'_0, y_0)\}$.[9] As shown in Observation 2, the matching $P_1 = \{(a_0, b_0), (x_0, y'_0), (x'_0, y_0)\}$ is more popular than N_1.

The matchings P_2 and P_3 are marked in red and blue respectively in Fig. 4. It is easy to check that neither P_2 nor P_3 is more popular than N_1. So $M_1 = P_1$. Since $M_1 \subseteq M$, it follows that M is a popular matching in G that contains (a_0, b_0) and (x_0, y'_0). Since Theorem 9 holds in our instance G, it follows that ψ is 1-in-3 satisfiable.

(The $\Leftarrow$ direction.) Suppose ψ is 1-in-3 satisfiable. Since Theorem 9 holds in our instance G, we know there is a popular matching P in G that contains the edges (a_0, b_0) and (x_0, y'_0). So P has to also contain the edge (x'_0, y_0). We claim that $\Delta(P, N) > 0$.

Let us partition P into $P_0 \cup P_1$ where $P_1 = \{(a_0, b_0), (x_0, y'_0), (x'_0, y_0)\}$ and $P_0 = P \setminus P_1$. We have $\Delta(P, N) = \Delta(P_0, N_0) + \Delta(P_1, N_1)$. We have already checked that $\Delta(P_1, N_1) = 4 - 2 = 2$. Moreover, $\Delta(P_0, N_0) = 0$ by the popularity of P_0 and N_0 in G_0. So $\Delta(P, N) = 2$, i.e., the popular matching P defeats N. ◻

Lemma 7 shows that it is NP-hard to decide if there exists a popular matching that defeats a given matching N. This problem is NP-complete since a 'yes'-instance N has a popular matching (which is easy to verify [4,21]) that defeats it. Thus Theorem 3 stated in Section 1 follows.

[9] The matching $\{(a_0, z), (z', b_0), (x_0, y_0), (x'_0, y'_0)\}$ is not popular in G_1 since N_1 is more popular.

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 N can be considered to be *very* unpopular. So there is no 'viable alternative' that defeats a fairly popular matching. Hence fairly popular matchings are a meaningful generalization of popular 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 matchings that belong to the support of popular mixed 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.

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: properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89:3–44, 1998.
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. M. Bóna. *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory (Fourth Edition)*. World Scientific, New Jersey, 2017.
6. F. Brandl and T. Kavitha. Two problems in max-size popular matchings. *Algorithmica*, 81(7):2738–2764, 2019.
7. Canadian Resident Matching Service. How the matching algorithm works. http://carms.ca/algorithm.htm.
8. V. Chvátal. *Linear programming*. W. H. Freeman, New York, 1983.
9. A. Cseh. Popular matchings. Trends in Computational Social Choice, Ulle Endriss (ed.), 2017.
10. A. Cseh and T. Kavitha. Popular edges and dominant matchings. *Mathematical Programming*, 172(1):209–229, 2018.
11. L. Ehlers and T. Morrill. (Il)legal assignments in school choice. *The Review of Economic Studies*, 87(4):1837–1875, 2020.
12. Y. Faenza and T. Kavitha. Quasi-popular matchings, optimality, and extended formulations. *Mathematics of Operations Research*, 47(1):427–457, 2022.
13. 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.
14. T. Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45(2):233–284, 1992.
15. T. Feder. Network flow and 2-satisfiability. *Algorithmica*, 11(3):291–319, 1994.
16. T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28(1):103–126, 2003.
17. D. Gale and L. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.
18. D. Gale and M. Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11(3):223–232, 1985.
19. P. Gärdenfors. Match making: assignments based on bilateral preferences. *Behavioural Science*, 20:166–173, 1975.
20. D. Gusfield and R. W. Irving. *The stable marriage problem: Structure and algorithms*. MIT Press, 1989.

21. C.-C. Huang and T. Kavitha. Popular matchings in the stable marriage problem. *Information and Computation*, 222:180–194, 2013.
22. C.-C. Huang and T. Kavitha. Popularity, mixed matchings, and self-duality. *Mathematics of Operations Research*, 46(2):405–427, 2021.
23. 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.
24. T. Kavitha. A size-popularity tradeoff in the stable marriage problem. *SIAM Journal on Computing*, 43(1):52–71, 2014.
25. 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.
26. 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.
27. 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.
28. T. Kavitha. Fairly popular matchings and optimality. In *Proceedings of the 39th International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 41:1–41:22, 2022.
29. T. Kavitha. Maximum matchings and popularity. *SIAM Journal on Discrete Mathematics*, 38(2):1202–1221, 2024.
30. T. Kavitha, J. Mestre, and M. Nasre. Popular mixed matchings. *Theoretical Computer Science*, 412:2679–2690, 2011.
31. 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.
32. National Resident Matching Program. Why the Match? http://www.nrmp.org/whythematch.pdf.
33. U. G. Rothblum. Characterization of stable matchings as extreme points of a polytope. *Mathematical Programming*, 54:57–67, 1992.
34. 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.
35. C.-P. Teo and J. Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23(4):874–891, 1998.
36. E. G. Thurber. Concerning the maximum number of stable matchings in the stable marriage problem. *Discrete Mathematics*, 248(1-3):195–219, 2002.
37. J. H. V. Vate. Linear programming brings marital bliss. *Operations Research Letters*, 8(3):147–153, 1989.

Appendix: Some Missing Details from Section 3.4

Stable matchings in multigraphs. Let $G = (A \cup B, E)$ be a multigraph where every vertex $u \in A \cup B$ has a strict preference order on its incident edges. It is well-known that all stable matchings in a simple graph match the same subset of vertices [18]. This property holds for stable matchings in multigraphs as well and it can be shown by LP duality (complementary slackness).

Proposition 3. *Every stable matching in the multigraph G matches the same subset of vertices.*

Proof. Let M be a stable matching in G. As done in Section 2.2, we will augment the edge set E with self-loops; so $\tilde{E} = E \cup \{(u,u) : u \in A \cup B\}$. Recall the function wt_M defined in Section 2.2: we have $\mathsf{wt}_M(u,u) \in \{0,-1\}$ for all $u \in A \cup B$ and $\mathsf{wt}_M(e) \in \{0, \pm 2\}$ for all $e \in E$. Furthermore, $\mathsf{wt}_M(e) \le 0$ for all $e \in E$ since M is stable (so it has no blocking edge). Consider (LP5).

$$\text{maximize} \sum_{e \in \tilde{E}} \mathsf{wt}_M(e) \cdot x_e \tag{LP5}$$

subject to

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

Recall that the optimal value of (LP5) is $\max_N \Delta(N,M)$ where the maximum is over all matchings N in G. Since M is stable, $\Delta(N,M) \leq 0$ for all N. The incidence vector of any stable matching S augmented with self-loops at unmatched vertices is an optimal solution to (LP5) since $\Delta(S,M) = 0$.

The linear program (LP6) is the dual LP. Since $\mathsf{wt}_M(e) \leq 0$ for all $e \in \tilde{E}$, $y_v = 0$ for all $v \in A\cup B$ is a feasible solution to (LP6). In fact, $\vec{y} = \vec{0}$ is an optimal solution to (LP6) since the optimal value of (LP6) is 0 (by LP duality).

$$\text{minimize} \sum_{v\in A\cup B} y_v \tag{LP6}$$

subject to

$$y_a + y_b \;\geq\; \mathsf{wt}_M(a,b) \;\; \forall\,(a,b)\in E \qquad \text{and} \qquad y_v \;\geq\; \mathsf{wt}_M(v,v) \;\; \forall\, v\in A\cup B.$$

For every vertex v matched in M, the constraint $y_v \geq \mathsf{wt}_M(v,v)$ is *slack* since $y_v = 0$ and $\mathsf{wt}_M(v,v) = -1$. It follows from complementary slackness that no optimal solution to (LP5) can contain the self-loop (v,v) where v is a vertex matched in M. In other words, for any stable matching S in G, $\{$vertices matched in $M\} \subseteq \{$vertices matched in $S\}$. By swapping the roles of M and S in the above argument, we have $\{$vertices matched in $S\} \subseteq \{$vertices matched in $M\}$. Hence the set of vertices matched in any two stable matchings in G is the same. ◻

Max-size popular matchings in G_c. We will show the following claim.

Claim 6 *Any vertex left unmatched in the max-size popular matching P_c in $G_c = (A_c \cup B_c, E_c)$ is an isolated vertex in G_c.*

Proof. Recall the graph H^* from Section 2.2. Analogous to how H^* was defined with respect to H, consider the graph G_c^* with respect to G_c. The graph G_c^0 is a more compact version of G_c^*: the difference between the graphs G_c^0 and G_c^* is that G_c^* contains dummy vertices, but G_c^0 has no dummy vertices. Recall that G_c^0 is a multigraph while G_c^* is a simple graph.

There is a natural bijection f between the set of stable matchings in G_c^* and the set of stable matchings in G_c^0. For any stable matching S in $G_c^* = (A_c^* \cup B_c^*, E_c^*)$:

$$\text{let } f(S) = \bigcup_{u\in A_c^*} \left(\{(u,v)_0 : (u,v)\in S\} \cup \{(u,v)_1 : (u',v)\in S\}\right).$$

It is straightforward to check that S is a stable matching in G_c^* if and only if $f(S)$ is a stable matching in G_c^0. A max-size popular matching in G_c can be computed by the *2-level Gale-Shapley* algorithm from [24]. It is known that running the 2-level Gale-Shapley algorithm in G_c is the same as running the Gale-Shapley algorithm in G_c^* [10]. By the equivalence between stable matchings in G_c^* and in G_c^0, running the Gale-Shapley algorithm in G_c^* is equivalent to running the Gale-Shapley algorithm in G_c^0.

Let S_c^0 be the stable matching computed by the Gale-Shapley algorithm in G_c^0 and let P_c be its *colorless* version. So P_c is a max-size popular matching in G_c. It is easy to prove the popularity of P_c via the following dual certificate $\vec{y}$. For each edge $(a,b) \in P_c$:

– if $(a,b)_0 \in S_c^0$: then let $y_a = 1$ and $y_b = -1$.
– if $(a,b)_1 \in S_c^0$: then let $y_a = -1$ and $y_b = 1$.

Also, $y_u = 0$ for every vertex u unmatched in P_c.

Thus $\sum_{v \in A_c \cup B_c} y_v = 0$. It is straightforward to check that $\vec{y}$ satisfies the constraints of the dual LP, i.e., (LP2) where wt_M is replaced by wt_{P_c} (see [6] for a proof of dual feasibility of $\vec{y}$). Thus $\vec{y}$ is an optimal solution to the dual LP. Moreover, the constraints corresponding to all edges incident to unmatched vertices are *slack*, i.e., for any vertex u unmatched in P_c and any neighbor v of u in G_c, we have $y_u = 0, y_v = 1$, and $\mathsf{wt}_{P_c}(u,v) = 0$, i.e., $y_u + y_v = 1 > 0 = \mathsf{wt}_{P_c}(u,v)$.

By complementary slackness, any fractional matching $\vec{q}$ that uses a *slack* edge (u,v) cannot be an optimal solution to the primal LP, i.e., $\mathsf{wt}_{P_c}(\vec{q}) = \Delta(\vec{q}, P_c) < 0$. In other words, P_c is more popular than $\vec{q}$. Thus no popular fractional edge of G_c is incident to any vertex left unmatched in P_c. Since every edge of G_c is a popular fractional edge in G (and so in G_c), this means any vertex left unmatched in P_c is an isolated vertex in G_c. □