Min-Cost Popular Matchings

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

Abstract

Let $G = (A \cup B, E)$ be a bipartite graph on n vertices where every vertex ranks its neighbors in a strict order of preference. A matching M in G is *popular* if there is *no* matching N such that vertices that prefer N to M outnumber those that prefer M to N. Popular matchings always exist in G since every stable matching is popular. Thus it is easy to find a popular matching in G – however it is NP-hard to compute a *min-cost* popular matching in G when there is a cost function on the edge set; moreover it is NP-hard to approximate this to any multiplicative factor. An $O^*(2^n)$ algorithm to compute a min-cost popular matching in G follows from known results. Here we show:

- an algorithm with running time $O^*(2^{n/4}) \approx O^*(1.19^n)$ to compute a min-cost popular matching;
- assume all edge costs are non-negative – then given $\varepsilon > 0$, a randomized algorithm with running time $\mathsf{poly}(n, \frac{1}{\varepsilon})$ to compute a matching M such that $\mathsf{cost}(M)$ is at most twice the optimal cost and with high probability, the fraction of all matchings more popular than M is at most $\frac{1}{2} + \varepsilon$.

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

Keywords and phrases Bipartite graphs, Stable matchings, Dual certificates

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2020.25

Funding We acknowledge support of the Department of Atomic Energy, Government of India, under project no. RTI4001.

Acknowledgements Thanks to Jannik Matuschke for conversations on semi-popular matchings and to Piyush Srivastava for helpful discussions on sampling matchings. Thanks to the reviewers for their helpful comments.

1 Introduction

Consider a matching problem in a bipartite graph $G = (A \cup B, E)$ on n vertices where every vertex has a strict ranking of its neighbors. Matching M is stable if M admits no *blocking edge* – an edge (a, b) is a blocking edge to M if a and b prefer each other to their respective assignments in M. Stable matchings always exist in G and one such matching can be computed in linear time by the classical Gale-Shapley algorithm [14]. Suppose there is a cost function on the edge set E. Computing a min-cost stable matching in G is a well-studied problem and there are several polynomial time algorithms to compute a min-cost stable matching and special variants of this problem [10, 11, 12, 21, 29, 30, 31].

Stability or absence of blocking edges is a rather strict notion – it is known that all stable matchings have the same size and match the same subset of vertices [15]. Consider the instance $G = (A \cup B, E)$ where $A = \{a_1, a_2\}$, $B = \{b_1, b_2\}$, and $E = \{(a_1, b_1), (a_1, b_2), (a_2, b_1)\}$. Suppose a_1 prefers b_1 to b_2 and similarly, b_1 prefers a_1 to a_2. The only stable matching here is $\{(a_1, b_1)\}$ whose size is half the size of the perfect matching $\{(a_1, b_2), (a_2, b_1)\}$.

A relaxation. In applications such as matching students to advisers, we would like to replace the notion of "no blocking edges" with a more relaxed notion of stability for the sake of obtaining a larger matching, or more generally, a more optimal matching. A natural relaxation of stability is the notion of *popularity* introduced by Gärdenfors [16] in 1975.

40th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2020).
Editors: Nitin Saxena and Sunil Simon; Article No. 25; pp. 25:1–25:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Roughly speaking, a matching is popular if there is no matching that makes more vertices happier. More formally, we say a vertex $u \in A \cup B$ prefers matching M to matching N if either (i) u is matched in M and unmatched in N or (ii) u is matched in both M and N and u prefers its partner in M to its partner in N. For any two matchings M and N, let $\phi(M, N)$ be the number of vertices that prefer M to N.

▶ **Definition 1.** *A matching M is* popular *if $\phi(M, N) \geq \phi(N, M)$ for every matching N in G, i.e., $\Delta(M, N) \geq 0$ where $\Delta(M, N) = \phi(M, N) - \phi(N, M)$.*

In an election between M and N where vertices cast votes, $\phi(M, N)$ is the number of votes for M and $\phi(N, M)$ is the number of votes for N. A popular matching never loses an election against another matching: thus it is a weak *Condorcet winner* [3, 4] in the corresponding voting instance. Although (weak) Condorcet winners need not exist in a general voting instance, popular matchings always exist in a bipartite graph since every stable matching is popular [16]. In fact, a stable matching is a min-size popular matching [19]. In the example described earlier, the perfect matching $\{(a_1, b_2), (a_2, b_1)\}$ is unstable but popular.

Efficient algorithms are known to compute a max-size popular matching in G [19, 23]. Though computing a min-size/max-size popular matching is easy, surprisingly, it is NP-hard to decide if G admits a popular matching that is not a min-size/max-size popular matching [9]. Also, computing a min-cost popular matching is NP-hard [9] when there is a cost function on the edge set. The min-cost popular matching problem includes other optimization problems such as computing a popular matching with forced/forbidden edges or one with max-utility as special cases and these variants are also NP-hard [9].

In applications such as matching students to advisers or medical residents to hospitals, where matchings have a long-term impact, it may be worthwhile to spend (exponential) time and compute an optimal popular matching in G. It follows from recent work on finding popular matchings in non-bipartite graphs [25] that there is an $O^*(2^n)$ time algorithm for the min-cost popular matching problem in a bipartite graph on n vertices (note that $O^*(2^n)$ stands for $O(2^n \cdot \mathsf{poly}(n))$). Here we study faster exponential time algorithms for this problem and show the following result.

▶ **Theorem 2.** *Given a bipartite graph $G = (A \cup B, E)$ on n vertices where every vertex has a strict preference list ranking its neighbors and a function $\mathsf{cost} : E \to \mathbb{R}$, a min-cost popular matching in G can be computed in $O^*(2^{n/4}) \approx O^*(1.19^n)$ time.*

The running time of our algorithm is $O(2^p \cdot \mathsf{poly}(n))$ where p is the number of connected components of size at least 4 in a special subgraph of G. Thus our algorithm is an FPT algorithm parameterized by p and when $p = O(\log n)$, this is a polynomial time algorithm.

When edge costs are non-negative, the *max-cost* popular matching problem in G admits an efficient 1/2-approximation algorithm – however the *min-cost* popular matching problem is NP-hard to approximate within any multiplicative factor even when edge costs are in $\{0, 1\}$ [9]. This motivates the following question: when edge costs are non-negative, is there an efficient algorithm to compute an *approximately popular* matching whose cost is $O(\mathsf{opt})$, where opt is the cost of a min-cost popular matching?

There are several ways to define an approximately popular matching and we choose the following novel definition: a matching M such that $\phi(M, N) \geq \phi(N, M)$ for a majority of matchings N in G. This motivates the definition of a *semi-popular* matching as follows.

▶ **Definition 3.** *Call a matching M in $G = (A \cup B, E)$* semi-popular *if $\phi(M, N) \geq \phi(N, M)$ for at least half the matchings N in G.*

Though semi-popularity is a natural relaxation of popularity, the set of semi-popular matchings seems to lack the structure of the set of popular matchings. We do not know how to efficiently test if a given matching is semi-popular or not. We show the following result on computing an *almost semi-popular* matching in an instance G with non-negative edge costs.

▶ **Theorem 4.** *Given a bipartite graph $G = (A \cup B, E)$ with $\mathsf{cost} : E \to \mathbb{R}_{\geq 0}$ and $\varepsilon > 0$, a matching M can be computed in $\mathsf{poly}(n, \frac{1}{\varepsilon})$ time such that $\mathsf{cost}(M) \leq 2\mathsf{opt}$ and with high probability M is undefeated by at least $1/2 - \varepsilon$ fraction of all matchings in G.*

Using the notation of bi-criteria approximation algorithms (see [27]), the above result is with high probability a $(2, \frac{1}{2} - \varepsilon)$ approximation of a min-cost popular matching, where the first coordinate is the ratio of the cost of our matching and opt and the second coordinate is a measure of popularity of our matching, more precisely, it is the fraction of matchings in G that our matching does not lose to. Designing an efficient algorithm to compute an $(O(1), 1 - \varepsilon)$ bi-criteria approximation is an open problem.

1.1 Background and Related Results

Algorithmic questions in the domain of popular matchings have been studied in the last 10-15 years. We refer to [5] for a survey. Initially, algorithms for popular matchings in instances with *one-sided* preferences (only vertices in A have preferences) were studied [1]. In the domain of two-sided preferences with ties, it is NP-complete to decide if popular matchings exist or not [2, 6]. The problem of deciding if a non-bipartite graph with strict preferences admits a popular matching is NP-complete [9, 17]. Popular matchings always exist in bipartite graphs with strict preferences [16]. However, as mentioned earlier, it is NP-hard to compute or approximate a min-cost popular matching. In order to cope with this hardness of approximation, a relaxation of popularity called *quasi-popularity* was considered in [8].

A matching M is quasi-popular if $\phi(N, M) \leq 2 \cdot \phi(M, N)$ for all matchings N. That is, M may lose many elections, however the *factor* of defeat, i.e., the ratio of number of votes won by the rival matching and the number of votes won by M, is bounded by 2. On the other hand, a semi-popular matching does not lose too many elections. A polynomial time algorithm to compute a quasi-popular matching of cost at most opt was given in [8].

There is a vast literature on fast exponential time algorithms for NP-hard problems and we refer to the book [13] on this subject. An algorithm with running time $O^*(c^n)$, where $c = O(1)$, was given in [25] to decide if a non-bipartite graph on n vertices with strict preferences has a popular matching or not. Fast exponential time algorithms for other hard problems in matchings under preferences are also known, e.g., the sex-equal stable marriage problem in bipartite graphs where the objective is to find a *fair* stable matching – a fast exponential time algorithm is known for this problem when the length of preference lists of vertices on one side of the bipartite graph is bounded from above by a small value [28].

1.2 Techniques

An $O^*(2^n)$ time algorithm was given in [25] to decide if a special popular matching called a *truly popular* matching exists or not in a general graph (not necessarily bipartite) on n vertices. A truly popular matching is a matching that is *popular fractional* (defined in Section 4). In bipartite graphs, every popular matching is truly popular and so this algorithm leads to an algorithm with running time $O^*(2^n)$ to compute a min-cost popular matching in the bipartite graph G. Our faster exponential time algorithm is an extension of this algorithm.

The earlier algorithm. The $O^*(2^n)$ time algorithm uses dual certificates or *witnesses* for popular matchings, where a witness $\vec{\alpha}$ is a vector in $\{0, \pm 1\}^n$ that obeys certain constraints (see Theorem 5). Corresponding to each of the 2^n parity combinations – whether α_u is 0 or ± 1 for each vertex u – the $O^*(2^n)$ algorithm constructs a stable matching instance and shows that every stable matching in this instance that avoids certain edges maps to a popular matching in G. Conversely, every popular matching in G can be realized as a stable matching that avoids certain edges in one of these 2^n instances. Computing a min-cost stable matching that excludes certain edges in each of these 2^n instances and taking the least cost such matching leads us to a min-cost popular matching in G.

Our faster algorithm. It was shown in [8] that all vertices in the same connected component in a subgraph G_0 of G called its "popular subgraph" have the same parity of their α-values. So instead of considering individual vertices, we consider non-trivial connected components in G_0 as our "units". Our main idea is that it suffices for the algorithm to go through parity combinations of α-values only for connected components in G_0 of size at least 4. So our algorithm constructs at most $2^{n/4}$ stable matching instances. However our stable matching instances are more elaborate than in the earlier algorithm and the most technical part of the analysis is the proof that stable matchings that avoid certain edges in such an instance map to popular matchings in G. The algorithm and its proof of correctness are given in Section 3.

Our bi-criteria approximation algorithm. Unlike the popular matching polytope, the popular fractional matching polytope has a compact extended formulation [26]. Thus a min-cost popular fractional matching can be computed in polynomial time by linear programming over this polytope. It is known that this polytope is half-integral [20]. Thus we can efficiently find *two* matchings M_1, M_2 in G such that $(I_{M_1} + I_{M_2})/2$ is a min-cost popular fractional matching in G, where I_M is the edge incidence vector of matching M. This implies that one of M_1, M_2 is semi-popular.

Interestingly, we do not know how to efficiently decide which of M_1, M_2 is semi-popular. We use the random sampler from [22] to sample matchings from a distribution close to the uniform distribution – this allows us to decide with high probability whether both M_1 and M_2 are almost semi-popular or one of them is not. This result is given in Section 4.

2 Popular Matchings and Witnesses

Let $\tilde{G}$ be the graph G augmented with self-loops. We assume that each vertex is its own last choice neighbor. Any matching M in G can henceforth be regarded as a perfect matching $\tilde{M}$ in $\tilde{G}$ by adding self-loops for all vertices left unmatched in M. The following edge weight function wt_M in $\tilde{G}$ will be useful to us. For any edge (a,b) in G, define:

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

So $\mathsf{wt}_M(e) = 0$ for every edge $e \in M$. We need to define wt_M on self-loops also. For any vertex $u \in A \cup B$, let $\mathsf{wt}_M(u,u) = 0$ if $(u,u) \in \tilde{M}$, else $\mathsf{wt}_M(u,u) = -1$. Let $\tilde{E} = E \cup \{(u,u) : u \in A \cup B\}$. For any matching N in G, we have:

$$\mathsf{wt}_M(\tilde{N}) = \sum_{e \in \tilde{N}} \mathsf{wt}_M(e) = \phi(N,M) - \phi(M,N) = \Delta(N,M).$$

Hence M is popular in G if and only if every perfect matching in the graph $\tilde{G}$ (with edge weights given by wt_M) has weight at most 0. Consider the max-weight perfect matching LP in the graph $\tilde{G}$: this is (LP1) given below in variables x_e for $e \in \tilde{E}$. Here $\tilde{\delta}(u) = \delta(u) \cup \{(u,u)\}$ for $u \in A \cup B$. The linear program (LP2) in variables α_u for $u \in A \cup B$ is the dual LP.

$$\begin{aligned} \max \sum_{e \in \tilde{E}} \mathsf{wt}_M(e) \cdot x_e \qquad & & \text{(LP1)} \\ \text{s.t.} \quad \sum_{e \in \tilde{\delta}(u)} x_e = 1 \quad & \forall\, u \in A \cup B \\ x_e \geq 0 \quad & \forall\, e \in \tilde{E}. \end{aligned}$$

$$\begin{aligned} \min \sum_{u \in V} \alpha_u \qquad & & \text{(LP2)} \\ \text{s.t.} \quad \alpha_a + \alpha_b \geq \mathsf{wt}_M(a,b) \quad & \forall\, (a,b) \in E \\ \alpha_u \geq \mathsf{wt}_M(u,u) \quad & \forall\, u \in A \cup B \end{aligned}$$

The characterization of popular matchings given in Theorem 5 follows from LP-duality and total unimodularity of the system. Recall that $|A \cup B| = n$.

▶ **Theorem 5** ([24, 26]). *A matching M in $G = (A \cup B, E)$ is popular if and only if there exists a vector $\vec{\alpha} \in \{0, \pm 1\}^n$ such that $\sum_{u \in A \cup B} \alpha_u = 0$,*

$$\alpha_a + \alpha_b \;\; \geq \;\; \mathsf{wt}_M(a,b) \quad \forall\, (a,b) \in E \quad \textit{and} \quad \alpha_u \;\; \geq \;\; \mathsf{wt}_M(u,u) \quad \forall\, u \in A \cup B.$$

Proof. The linear program (LP2) admits an optimal solution that is integral since its constraint matrix is totally unimodular. The vector $\vec{\alpha}$ is an integral optimal solution of (LP2). We have $\alpha_u \geq \mathsf{wt}_M(u,u) \geq -1$ for all u.

Since $\tilde{M}$ is an optimal solution to (LP1), complementary slackness implies $\alpha_u + \alpha_v = \mathsf{wt}_M(u,v) = 0$ for each edge $(u,v) \in M$. Thus $\alpha_u = -\alpha_v \leq 1$ for every vertex u matched to a non-trivial neighbor v in M. Regarding any vertex u such that $(u,u) \in \tilde{M}$, we have $\alpha_u = \mathsf{wt}_M(u,u) = 0$ (by complementary slackness). Hence $\vec{\alpha} \in \{0, \pm 1\}^n$. ◀

▶ **Definition 6.** *For any popular matching M, a vector $\vec{\alpha} \in \{0, \pm 1\}^n$ as given in Theorem 5 is called a* witness *of M.*

A popular matching may have several witnesses. A stable matching S has $\vec{0}$ as a witness, since $\mathsf{wt}_S(e) \leq 0$ for all edges e in $\tilde{G}$. Call an edge e in $G = (A \cup B, E)$ *popular* if there is some popular matching in G that contains e. Let E_0 be the set of popular edges in G. The set E_0 can be computed in linear time [7]. Call the subgraph $G_0 = (A \cup B, E_0)$ the *popular subgraph* of G. The following property will be very useful.

▶ **Lemma 7** ([8]). *Let M be any popular matching in G and let $\vec{\alpha}$ be any witness of M. In any connected component C in the popular subgraph G_0: either (i) $\alpha_u = 0$ for all $u \in C$ or (ii) $\alpha_u \in \{\pm 1\}$ for all $u \in C$.*

Proof. Consider any popular edge (a,b). So there is some popular matching N that contains (a,b). Since $\mathsf{wt}_M(\tilde{N}) = \Delta(N,M) = 0$ (because M and N are popular matchings), $\tilde{N}$ is an optimal solution to (LP1). We know that $\vec{\alpha}$ is an optimal solution to (LP2). So it follows from complementary slackness that $\alpha_a + \alpha_b = \mathsf{wt}_M(a,b)$. Since $\mathsf{wt}_M(a,b) \in \{\pm 2, 0\}$ (an even number), the integers α_a and α_b have the same parity.

Let u and v be any 2 vertices in the same connected component in the popular subgraph G_0. So there is a u-v path ρ in G such that every edge in ρ is a popular edge. We have just seen that the endpoints of each popular edge have the same parity in $\vec{\alpha}$. Hence α_u and α_v have the same parity. Thus either $\alpha_u = 0$ for all $u \in C$ or $\alpha_u \in \{\pm 1\}$ for all $u \in C$. ◀

3 A fast exponential time algorithm for min-cost popular matching

Let $C_1, \ldots, C_r$ be the connected components in the popular subgraph G_0. Assume $C_1, \ldots, C_q$ are the non-trivial components, i.e., $|C_i| \geq 2$ for $1 \leq i \leq q$ and $|C_i| = 1$ for $q+1 \leq i \leq r$. So each of $C_{q+1}, \ldots, C_r$ consists of a single vertex that is left unmatched in all popular matchings in G. Call such a vertex *unpopular*. Let U be the set of unpopular vertices. The following two observations will be useful.

▶ **Observation 1.** *Let M be a popular matching with $\vec{\alpha}$ as a witness. If $u \in U$ then $\alpha_u = 0$.*

Proof. Since M leaves u unmatched, the self-loop $(u, u) \in \tilde{M}$. Observe that $\tilde{M}$ is an optimal solution to (LP1) and $\vec{\alpha}$ is an optimal solution to (LP2). So $\alpha_u = \mathsf{wt}_M(u, u) = 0$ by complementary slackness. ◁

▶ **Observation 2.** *Every non-trivial component C in the popular subgraph G_0 has an even number of vertices.*

Proof. All max-size popular matchings in G leave the same vertices unmatched and these unmatched vertices are unpopular [18]. Thus a max-size popular matching M restricted to every non-trivial component C in G_0 is *perfect*, i.e., all vertices in C are matched in M. Hence $|C|$ is even. ◁

Let $C_1, \ldots, C_p$ be the components in G_0 of size greater than 2. This means $|C_i| \geq 4$ for $i \in [p]$ (by Observation 2). So $C_{p+1}, \ldots, C_q$ are the components in G_0 of size exactly 2.

For every subset $I \subseteq \{1, \ldots, p\}$, our algorithm builds a corresponding graph G_I. Among all stable matchings in G_I that satisfy certain constraints, our algorithm finds a min-cost matching (call it N_I). It will be shown that among all subsets $I \subseteq [p]$, the matching N_I with the least cost will map to a min-cost popular matching in G.

The new instance G_I. Let $I \subseteq [p]$. Partition the vertices in $A \cup B$ into three subsets:

$$S_0 = \cup_{i \in I} C_i \cup U, \qquad S_1 = \cup_{i \in [p] \setminus I} C_i, \qquad \text{and} \qquad S_2 = \cup_{i=p+1}^{q} C_i.$$

Our goal is to build G_I such that all popular matchings in G that admit witnesses $\vec{\alpha}$ where $\alpha_u = 0$ for $u \in S_0$ and $\alpha_u \in \{\pm 1\}$ for $u \in S_1$ become *stable matchings* in G_I. For vertices in S_2, we do not a priori commit any particular α-value. This is reflected in the vertex set V_I:

$$V_I = \{u_0 : u \in S_0 \cup S_2\} \,\cup\, \{u_+, u_-, d(u) : u \in S_1 \cup S_2\} \,\cup\, \{d'(u) : u \in S_2\}.$$

The set V_I contains a single vertex u_0 for every $u \in S_0$, three vertices $u_+, u_-, d(u)$ for every $u \in S_1$, and five vertices $u_+, u_-, u_0, d(u), d'(u)$ for every $u \in S_2$. Since the α-value of every $u \in S_0$ is fixed to be 0, we have a unique vertex u_0 in G_I for each $u \in S_0$.

Since the α-value of every $u \in S_1$ is either 1 or -1, there are two vertices u_+, u_- in G_I for each $u \in S_1$. However in order to map stable matchings in G_I to matchings in G, we want at most one of u_+, u_- to be matched in any stable matching in G_I: this is achieved by using a *dummy vertex* $d(u)$. Preferences will be such that one of u_+, u_- has to be matched to $d(u)$ in any stable matching in G_I. So every stable matching in G_I matches at most one of u_+, u_- to a non-dummy neighbor.

Since the α-value of every $u \in S_2$ is one of $0, \pm 1$, we have *three* vertices u_+, u_-, u_0 in G_I for each $u \in S_2$. However we want at most one of u_+, u_-, u_0 to be matched in any stable matching in G_I and this is achieved by using *two* dummy vertices $d(u)$ and $d'(u)$.

Preferences will be such that two of u_+, u_-, u_0 have to be matched to $d(u)$ and $d'(u)$ in any stable matching in G_I. So every stable matching in G_I matches at most one of u_+, u_-, u_0 to a non-dummy neighbor.

The edge set E_I of the instance G_I is defined as follows. For every $(u, v) \in E$, the edge set E_I consists of one or more of the following edges: (i) (u_0, v_0), (ii) (u_+, v_0), (iii) (u_0, v_+), (iv) (u_-, v_+), (v) (u_+, v_-).

In more detail, let $u \in A \cup B$. Let v be a neighbor of u in G.

- if $u, v \in S_0$ then (u_0, v_0) is in E_I.
- if $u \in S_1$ and $v \in S_0$ then (u_+, v_0) is in E_I.
- if $u, v \in S_1$ and u prefers v to every neighbor in S_0 then (u_-, v_+) is in E_I.

The edges in G_I that correspond to edges (u, v) in G with an endpoint, say $u \in A \cup B$, in S_2 are described below.

- let $v \in S_0$. If u prefers v to its "popular partner"[1] then the edge $(u_+, v_0) \in E_I$; else the edge $(u_0, v_0) \in E_I$.
- let $v \in S_1$. If u prefers v to its popular partner then the edge $(u_0, v_+) \in E_I$. If v prefers u to every neighbor in S_0 then the edge $(u_+, v_-) \in E_I$.
- let $v \in S_2$. If either v is u's popular partner or one of u, v prefers the other to its popular partner[2] then the edge $(u_0, v_0) \in E_I$. Moreover, if u prefers v to every neighbor in S_0 then the edge $(u_-, v_+) \in E_I$.

For every $u \in S_1$: the edges $(u_+, d(u))$ and $(u_-, d(u))$ are in E_I. For every $u \in S_2$: the edges $(u_+, d(u)), (u_0, d(u))$ and the edges $(u_0, d'(u)), (u_-, d'(u))$ are in E_I.

Vertex preferences. We will first list preference orders for dummy vertices.

- For $u \in S_1$: $d(u)$'s preference order is $u_+ \succ u_-$, i.e., top choice u_+ followed by u_-.
- For $u \in S_2$: $d(u)$'s preference order is $u_+ \succ u_0$ and $d'(u)$'s preference order is $u_0 \succ u_-$.

Let $u \in A \cup B$. We now list preference orders for u_+, u_0, and u_-. An observation that will be useful here is that for any two adjacent vertices u, v in G, there is at most one element in $\{v_0, v_+, v_-\}$ in the preference list of u_+; similarly, in the preference lists of u_0 and u_-.

1. For $u \in S_0$: u_0's preference order among its neighbors in G_I is as per u's preference order in G, i.e., ignore subscripts of vertices and arrange them as per u's preference order in G.
2. For $u \in S_1 \cup S_2$: u_+'s preference order among its neighbors in G_I is as per u's preference order in G with $d(u)$ as its least preferred neighbor.
3. For $u \in S_1$ (resp., $u \in S_2$): u_-'s preference order among its neighbors in G_I is $d(u)$ (resp., $d'(u)$) as its top choice neighbor followed by its other neighbors in G_I as per u's preference order in G.
4. For $u \in S_2$: u_0's order among its neighbors in G_I is $d(u)$ as its top choice neighbor followed by its other neighbors in G_I as per u's preference order in G and $d'(u)$ as its least preferred neighbor.

For $(a, b) \in E$ and $x, x' \in \{0, \pm\}$, for every $(a_x, b_{x'}) \in E_I$, we set $\mathsf{cost}(a_x, b_{x'}) = \mathsf{cost}(a, b)$. Also, the cost of any edge incident to a dummy vertex is 0.

[1] $u \in S_2$: so $u \in C_j$ where $|C_j| = 2$; hence all popular matchings in G match u to the same neighbor.
[2] Note that both u and v cannot prefer each other to their respective popular partners since that would make (u, v) a blocking edge to every stable matching in G.

▶ **Theorem 8.** *Let M be a popular matching in $G = (A \cup B, E)$ with a witness $\vec{\alpha} \in \{0, \pm 1\}^n$ where $\alpha_v = 0$ for $v \in S_0$ and $\alpha_v \in \{\pm 1\}$ for $v \in S_1$. Then there exists a stable matching N_I in G_I such that $\mathsf{cost}(N_I) = \mathsf{cost}(M)$ and the following three properties are satisfied:*
1. *N_I avoids all edges between a subscript $+$ vertex and a subscript 0 vertex,*
2. *N_I matches all subscript $-$ vertices, and*
3. *N_I includes $q - p$ edges from the set $\cup_{i=p+1}^{q} \{(a_+, b_-), (a_0, b_0), (a_-, b_+) : a, b \in C_i\}$.*

Proof. M is a popular matching in $G = (A \cup B, E)$ with a witness $\vec{\alpha} \in \{0, \pm 1\}^n$. For any $u \in A \cup B$, we will define $s_u = +/-/0$ corresponding to $\alpha_u = +1/-1/0$, respectively. That is, (i) $\alpha_u = 1$ implies $s_u = +$, (ii) $\alpha_u = -1$ implies $s_u = -$, and (iii) $\alpha_u = 0$ implies $s_u = 0$.

- For $u \in S_1$: if $s_u = +$ then let $t_u = -$ else let $t_u = +$.
- For $u \in S_2$: if $s_u = +$ then let $t_u = 0$ and $t'_u = -$; if $s_u = 0$ then let $t_u = +$ and $t'_u = -$; if $s_u = -$ then let $t_u = +$ and $t'_u = 0$.

Define the set N_I as follows:

$$N_I = \{(a_{s_a}, b_{s_b}) : (a, b) \in M\} \ \cup\ \{(u_{t_u}, d(u)) : u \in S_1 \cup S_2\} \ \cup\ \{(u_{t'_u}, d'(u)) : u \in S_2\}.$$

We need to show that $N_I \subseteq E_I$, i.e., for every $(a, b) \in M$, the edge (a_{s_a}, b_{s_b}) is present in G_I. Observe that $\tilde{M}$ and $\vec{\alpha}$ are optimal solutions of (LP1) and (LP2), respectively. It follows from complementary slackness that $\alpha_a + \alpha_b = \mathsf{wt}_M(a, b) = 0$ for every $(a, b) \in M$. Thus either $\alpha_a = \alpha_b = 0$ or $\{\alpha_a, \alpha_b\} = \{-1, 1\}$.

For every edge (a, b) in M where $\alpha_a = \alpha_b = 0$ (each such edge is in $(S_0 \times S_0) \cup (S_2 \times S_2)$), observe that the edge (a_0, b_0) is in G_I. In particular, if $(a, b) \in (S_2 \times S_2) \cap M$, then we have $C_i = \{a, b\}$ for some $i \in \{p + 1, \ldots, q\}$ and we always include the edge (a_0, b_0) in G_I.

Consider an edge (a, b) in M where α_a or α_b is -1 (each such edge is in $(S_1 \times S_1) \cup (S_2 \times S_2)$). Assume wlog that $\alpha_a = -1$. Since $\vec{\alpha}$ is a witness of M, for every neighbor $c \in S_0$ of a, we have $\mathsf{wt}_M(a, c) \leq \alpha_a + \alpha_c = -1 + 0 = -1$. This means $\mathsf{wt}_M(a, c) = -2$, i.e., a prefers its partner in M (this is b) to c. The constraint $\mathsf{wt}_M(a, c) = -2$ holds for every neighbor c of a that is in S_0. Hence it follows from the definition of the edge set of G_I that (a_-, b_+) is in G_I.

Thus every edge of N_I is present in G_I, hence N_I is a matching in G_I. We will now show that N_I obeys properties (1)-(3) given in the statement of the theorem.
1. For every edge $(a, b) \in M$, we have $\alpha_a + \alpha_b = \mathsf{wt}_M(a, b) = 0$ (by complementary slackness). Thus every edge in N_I that is not incident to any dummy vertex is of the type (a_+, b_-) or (a_0, b_0) or (a_-, b_+). Hence N_I avoids all edges between a subscript 0 vertex and a subscript $+$ vertex.
2. For any vertex u left unmatched in M, we have $\alpha_u = \mathsf{wt}_M(u, u) = 0$ (by complementary slackness). So $u \in S_0 \cup S_2$. Since every vertex in S_2 is matched to its popular partner in all popular matchings in G, the unmatched vertex $u \in S_0$. Thus for every $u \in (A \cup B) \setminus S_0$, we have $(u, v) \in M$ for some neighbor v: if $\alpha_u = -1$ then $(u_-, v_+) \in N_I$ else either $(u_-, d(u))$ or $(u_-, d'(u))$ is in N_I. Thus all subscript $-$ vertices are matched in N_I.
3. For every connected component $C_i = \{a, b\}$ in G_0, where $p + 1 \leq i \leq q$, we know that $(a, b) \in M$. Thus one of $(a_+, b_-), (a_0, b_0), (a_-, b_+)$ is in N_I. So N_I includes $q - p$ edges from the set $\cup_{i=p+1}^{q} \{(a_+, b_-), (a_0, b_0), (a_-, b_+) : a, b \in C_i\}$.

We will now show that N_I is a stable matching in G_I. For any $u \in A \cup B$, it is easy to see there is no blocking edge with a dummy vertex as an endpoint. This is because a dummy vertex has only two neighbors and when it is matched to its second choice neighbor, its top choice neighbor is matched to a more preferred neighbor.

Regarding edges in E_I that correspond to edges in E, note that E_I contains certain edges of the form (a_0, b_0), (a_+, b_0), (a_0, b_+), (a_+, b_-), (a_-, b_+) for $(a, b) \in E$. We now need to show that no such edge in E_I blocks N_I. Consider any $(a, b) \in E$.

1. *Both a and b are in S_0:* so $\alpha_a = \alpha_b = 0$. We need to show that (a_0, b_0) is not a blocking edge to N_I. Since $\mathsf{wt}_M(a, b) \le \alpha_a + \alpha_b = 0$, either $(a, b) \in M$ or (at least) one of a, b is matched in M to a more preferred neighbor. That is, either $(a_0, b_0) \in N_I$ or one of a_0, b_0 is matched in N_I to a more preferred neighbor. So (a_0, b_0) does not block N_I.
2. *One of a, b is in S_0 and the other is in S_1:* assume wlog that $a \in S_0$ and $b \in S_1$. So $\alpha_a = 0$ and $\alpha_b \in \{\pm 1\}$. We need to show that (a_0, b_+) is not a blocking edge to N_I. There are two subcases here: (i) $\alpha_b = 1$ and (ii) $\alpha_b = -1$.
   In the first subcase, $\mathsf{wt}_M(a, b) \le \alpha_a + \alpha_b = 1$ which implies $\mathsf{wt}_M(a, b) \le 0$. So one of a, b is matched in M to a more preferred neighbor. So one of a_0, b_+ is matched in N_I to a more preferred neighbor. Hence (a_0, b_+) does not block N_I.
   In the second subcase, $\mathsf{wt}_M(a, b) \le \alpha_a + \alpha_b = -1$ which implies $\mathsf{wt}_M(a, b) = -2$. So *both* a and b are matched in M to more preferred neighbors. In particular, a_0 is matched in N_I to a neighbor preferred to b_+. Hence (a_0, b_+) does not block N_I.
3. *Both a and b are in S_1:* so $\alpha_a, \alpha_b \in \{\pm 1\}$. We need to show that the edges (a_-, b_+) and (a_+, b_-) (whichever of these is in E_I) do not block N_I. If $\alpha_a = \alpha_b = 1$ then both a_- and b_- are matched to their top choice neighbors $d(a)$ and $d(b)$, respectively. So neither (a_-, b_+) nor (a_+, b_-) blocks N_I.
   If $\alpha_a = 1$ and $\alpha_b = -1$ then $\mathsf{wt}_M(a, b) \le 0$. So either $(a, b) \in M$ or one of a, b is matched in M to a more preferred neighbor in G. That is, either $(a_+, b_-) \in N_I$ or one of a_+, b_- is matched in N_I to a more preferred neighbor in G_I. Moreover, the edge (a_-, b_+) cannot block N_I since a_- is matched in N_I to its top choice neighbor $d(a)$. The subcase when $\alpha_a = -1$ and $\alpha_b = 1$ is symmetric.
   The last subcase is $\alpha_a = \alpha_b = -1$. So $\mathsf{wt}_M(a, b) = -2$. Hence *both* a and b are matched in M to more preferred neighbors, i.e., both a_- and b_- are matched in N_I to neighbors preferred to b_+ and a_+, respectively. So neither (a_-, b_+) nor (a_+, b_-) blocks N_I.

The proofs for the remaining three cases (when at least one of a, b is in S_2) are given below in Claims 9-11. Thus N_I is a stable matching in G_I. ◀

▷ Claim 9. Suppose one of a, b (say, b) is in S_0 and a is in S_2. Then neither (a_+, b_0) nor (a_0, b_0) blocks N_I.

Proof. Since $a \in S_2$ and $b \in S_0$, we have $\alpha_a \in \{0, \pm 1\}$ and $\alpha_b = 0$. Suppose $\alpha_a = -1$. Then $\mathsf{wt}_M(a, b) \le -1$, i.e., $\mathsf{wt}_M(a, b) = -2$. So both a and b are matched in M to more preferred neighbors. Since M always matches a to its *popular partner*, it means a prefers its popular partner to b. Thus (a_0, b_0) is in E_I and b_0 is matched in N_I to a neighbor preferred to a_0.

Suppose $\alpha_a \in \{0, 1\}$. Then $\mathsf{wt}_M(a, b) \le 1$, i.e., $\mathsf{wt}_M(a, b) \le 0$. So one of a, b is matched in M to a more preferred neighbor. Either (i) (a_0, b_0) is in E_I and so a_0 is matched in N_I to a more preferred neighbor (its popular partner or $d(a)$) than b_0 or (ii) (a_+, b_0) is in E_I, in which case a prefers b to its popular partner – so b has to be matched in M to a neighbor preferred to a, i.e., b_0 is matched in N_I to a neighbor preferred to a_+. Hence neither (a_+, b_0) nor (a_0, b_0) (whichever is present in E_I) blocks N_I. ◁

▷ Claim 10. Suppose one of a, b (say, b) is in S_1 and a is in S_2. Then neither (a_0, b_+) nor (a_+, b_-) blocks N_I.

Proof. Since $a \in S_2$ and $b \in S_1$, we have $\alpha_a \in \{0, \pm 1\}$ and $\alpha_b \in \{\pm 1\}$. Suppose a prefers b to its popular partner. Then (a_0, b_+) is in E_I and also $\mathsf{wt}_M(a, b) \ge 0$. If $\alpha_a = 1$ then a_0 is matched to its most preferred neighbor $d(a)$ and so (a_0, b_+) does not block N_I. If $\alpha_a \le 0$

then $\alpha_b = 1$ since $\alpha_a + \alpha_b \geq \mathsf{wt}_M(a,b) \geq 0$. Also $\mathsf{wt}_M(a,b) \leq 1$ since $\alpha_a + \alpha_b = 1$, i.e., $\mathsf{wt}_M(a,b) = 0$. So b has to be matched in M to a neighbor preferred to a, i.e., b_+ has to be matched in N_I to a neighbor preferred to a_0. Hence (a_0, b_+) does not block N_I.

Suppose b prefers a to all neighbors in S_0. Then (a_+, b_-) is in E_I. If $\alpha_b = 1$ then b_- is matched to its most preferred neighbor $d'(b)$ in N_I. Suppose $\alpha_b = -1$. If $\alpha_a \in \{0, -1\}$ then $\mathsf{wt}_M(a,b) \leq \alpha_a + \alpha_b \leq -1$. So $\mathsf{wt}_M(a,b) = -2$. This means both a, b are matched in M to more preferred neighbors. Hence b_- is matched in N_I to a neighbor preferred to a_+. Suppose $\alpha_a = 1$. Then $\mathsf{wt}_M(a,b) \leq 0$: so one of a, b is matched in M to a more preferred neighbor. So one of a_+, b_- is matched in N_I to a more preferred neighbor. Thus the edge (a_+, b_-) does not block N_I. ◁

▷ Claim 11. Suppose both a and b are in S_2. Then none of the edges $(a_0, b_0), (a_+, b_-), (a_-, b_+)$ blocks N_I.

Proof. Since a, b are in S_2, we have $\alpha_a, \alpha_b \in \{0, \pm 1\}$. If a, b are each other's popular partners or one of them prefers the other to its popular partner then the edge (a_0, b_0) is in E_I and also $\mathsf{wt}_M(a,b) \geq 0$. So either $\alpha_a = \alpha_b = 0$ or at least one of α_a, α_b is 1. So either $(a_0, b_0) \in N_I$ or one of a_0, b_0 is matched in N_I to a more preferred neighbor. Thus (a_0, b_0) does not block N_I.

If a prefers b to all its neighbors in S_0 then the edge (a_-, b_+) is in E_I. If $\alpha_a \in \{0, 1\}$ then a_- is matched to its most preferred neighbor $d'(a)$ in N_I. So the edge (a_-, b_+) does not block N_I. Suppose $\alpha_a = -1$. If $\alpha_b \in \{0, -1\}$ then $\mathsf{wt}_M(a,b) \leq \alpha_a + \alpha_b \leq -1$. So $\mathsf{wt}_M(a,b) = -2$. This means both a, b are matched in M to more preferred neighbors. Hence a_- is matched in N_I to a neighbor preferred to b_+. Suppose $\alpha_b = 1$. Then $\mathsf{wt}_M(a,b) \leq 0$: so one of a, b is matched in M to a more preferred neighbor. So one of a_-, b_+ is matched in N_I to a more preferred neighbor. Thus the edge (a_-, b_+) does not block N_I.

The analysis that (a_+, b_-) does not block N_I when b prefers a to all neighbors in S_0 is analogous. ◁

Let us call a stable matching in G_I that satisfies the three properties given in Theorem 8 a *desired* stable matching. Theorem 12 proves the converse of Theorem 8.

▶ **Theorem 12.** *Suppose G_I admits a desired stable matching, say N_I. Then N_I can be mapped to a popular matching M in G such that $\mathsf{cost}(N_I) = \mathsf{cost}(M)$.*

Proof. The matching M will be defined as follows:

$$M = \{(a,b) : (a_{s_a}, b_{s_b}) \in N_I \text{ for } s_a, s_b \in \{0, \pm\}\}.$$

For any $u \in A \cup B$, at most one of u_+, u_0, u_- can be matched to a non-dummy neighbor in N_I. Thus M is a valid matching in G. In order to prove M's popularity, we will show a witness $\vec{\alpha} \in \{0, \pm 1\}^n$. Define $\alpha_u = 0$ for all $u \in S_0$. Let $u \in S_1$. Since N_I is stable, the vertex $d(u)$ (as the top choice neighbor of u_-) has to be matched in N_I. So for $u \in S_1$, define α_u as follows:

$$\text{let } \alpha_u = \begin{cases} -1 & \text{if } (u_+, d(u)) \in N_I \\ 1 & \text{if } (u_-, d(u)) \in N_I. \end{cases}$$

Let $u \in S_2$. Then there are two dummy vertices $d(u)$ and $d'(u)$ for u and both of them (as the top choice neighbors of u_0 and u_-, resp.) have to be matched in N_I. So for $u \in S_2$, define α_u as follows:

$$\text{let } \alpha_u = \begin{cases} -1 & \text{if } (u_+, d(u)) \text{ and } (u_0, d'(u)) \text{ are in } N_I \\ 0 & \text{if } (u_+, d(u)) \text{ and } (u_-, d'(u)) \text{ are in } N_I \\ 1 & \text{if } (u_0, d(u)) \text{ and } (u_-, d'(u)) \text{ are in } N_I. \end{cases}$$

We will now show that $\vec{\alpha}$ is a witness of M's popularity. Observe that all edges in N_I not involving any dummy vertex are of the form (a_+, b_-) or (a_0, b_0) or (a_-, b_+). This is because N_I avoids all edges of the type (a_+, b_0) and (a_0, b_+) (by property (1)). Thus $\alpha_a + \alpha_b = 0$ for all $(a, b) \in M$. Due to property (2), property (3), and N_I's stability, it follows that for any vertex u left unmatched in M, we have $u \in S_0$, i.e., $\alpha_u = 0$. So $\sum_{u \in A \cup B} \alpha_u = 0$.

It is also easy to see that $\alpha_u \geq \mathsf{wt}_M(u, u)$ for every vertex u. This is because every vertex $u \in (A \cup B) \setminus S_0$ is matched in M and so we have $\alpha_u \geq -1 = \mathsf{wt}_M(u, u)$ for these vertices. For any vertex $u \in S_0$, we have $\alpha_u = 0 \geq \mathsf{wt}_M(u, u)$.

What is left to show is that every edge (a, b) in G is *covered*, i.e., $\alpha_a + \alpha_b \geq \mathsf{wt}_M(a, b)$. This is proved below in Lemma 13. Thus $\vec{\alpha}$ is a witness of M (by Theorem 5). So M is a popular matching; also $\mathsf{cost}(M) = \mathsf{cost}(N_I)$. This finishes the proof of Theorem 12. ◀

▶ **Lemma 13.** *We have $\alpha_a + \alpha_b \geq \mathsf{wt}_M(a, b)$ for every edge (a, b) in G.*

Proof. Recall that $\mathsf{wt}_M(a, b) \in \{0, \pm 2\}$. Any edge (a, b) where $\alpha_a = \alpha_b = 1$ is obviously covered since $\mathsf{wt}_M(a, b) \leq 2$. The proofs for other cases of (α_a, α_b) are given in Claims 14-18.

▷ Claim 14. Any edge (a, b) where $\{\alpha_a, \alpha_b\} = \{0, 1\}$ is covered.

Proof. Assume without loss of generality $\alpha_a = 1$ and $\alpha_b = 0$: so $a \in S_1 \cup S_2$ and $b \in S_0 \cup S_2$. If the edge (a_+, b_0) is in E_I then the stability of N_I implies that either (i) a_+ is matched in N_I to a neighbor preferred to b_0 or (ii) b_0 is matched in N_I to a neighbor preferred to a_+ (moreover, a non-dummy neighbor since $\alpha_b = 0$). So at least one of a, b is matched in M to a more preferred neighbor. Thus $\mathsf{wt}_M(a, b) \leq 0$.

The edge (a_+, b_0) is *not* present in G_I in the following 2 cases:

1. both a, b are in S_2 and either (i) a, b are each other's *popular partners* or (ii) at least one of a, b prefers its popular partner to the other (see footnote 2). By property (3), every vertex in S_2 is matched in M to its popular partner. Hence $\mathsf{wt}_M(a, b) \leq 0$.
2. either (i) $a \in S_2$ prefers its popular partner (call it y) to $b \in S_0$ or (ii) $b \in S_2$ prefers its popular partner (call it z) to $a \in S_1$; property (3) forces (a, y) to be in M in the first case and (z, b) to be in M in the second case. So $\mathsf{wt}_M(a, b) \leq 0$.

Hence in all cases, we have $\mathsf{wt}_M(a, b) \leq 0 < 1 = \alpha_a + \alpha_b$. ◁

▷ Claim 15. Any edge (a, b) where $\alpha_a = \alpha_b = 0$ is covered.

Proof. Since $\alpha_a = \alpha_b = 0$, we have $a, b \in S_0 \cup S_2$. If the edge (a_0, b_0) is in G_I, then it follows from the stability of N_I that $(a_0, b_0) \in N_I$ or one of a_0, b_0 is matched in N_I to a more preferred (non-dummy) neighbor, i.e., at least one of a, b is matched in M to a more preferred neighbor. Thus $\mathsf{wt}_M(a, b) \leq 0$.

The edge (a_0, b_0) is *not* present in G_I in the following 2 cases:

1. $a \in S_2$ prefers $b \in S_0$ to its popular partner: in this case (a_+, b_0) is in G_I. Since $\alpha_a = 0$, the vertex a_+ is matched in N_I to its least preferred neighbor $d(a)$. Thus it follows from the stability of N_I that b_0 is matched to a more preferred neighbor than a_+, so $\mathsf{wt}_M(a, b) \leq 0$. It is similar when $b \in S_2$ prefers $a \in S_0$ to its popular partner.
2. both a, b are in S_2 and they prefer their respective popular partners to each other: in this case $\mathsf{wt}_M(a, b) = -2$.

Hence in all cases, we have $\mathsf{wt}_M(a, b) \leq 0 = \alpha_a + \alpha_b$. ◁

▷ Claim 16. Any edge (a, b) where $\{\alpha_a, \alpha_b\} = \{-1, 1\}$ is covered.

Proof. Assume without loss of generality that $\alpha_a = 1$ and $\alpha_b = -1$. We need to show that $\mathsf{wt}_M(a, b) \leq 0$. Either (i) $(a_+, b_-) \in N_I$ or (ii) (a_+, y_-) and (z_+, b_-) are in N_I for some neighbors y, z of a, b, respectively. In case (i), $\mathsf{wt}_M(a, b) = 0$. In case (ii), we will consider 2 subcases.

1. Suppose $a, b \in S_1$ or $a, b \in S_2$ or $a \in S_2$ and $b \in S_1$. Since the edge (z_+, b_-) is in G_I, b prefers z to all its neighbors in S_0. Hence if b prefers a to z then the edge (a_+, b_-) has to be present in G_I. It follows from the stability of N_I that a_+ prefers y_- to b_-, i.e., a prefers y to b. Hence $\mathsf{wt}_M(a, b) \leq 0$.
2. The remaining case is when $a \in S_1$ and $b \in S_2$. So z is b's popular partner. If b prefers a to z then the edge (a_+, b_0) is present in G_I. Since b_0 is matched to its least preferred neighbor $d'(b)$ in N_I, the stability of N_I implies that a_+ prefers y_- to b_0, i.e., a prefers y to b. Hence $\mathsf{wt}_M(a, b) \leq 0$.

Hence in all cases, we have $\mathsf{wt}_M(a, b) \leq 0 = \alpha_a + \alpha_b$. ◁

▷ Claim 17. Any edge (a, b) where $\alpha_a = \alpha_b = -1$ is covered.

Proof. So (a_-, y_+) and (z_+, b_-) are in N_I for some neighbors y, z of a, b, respectively. There are 3 cases here:

1. Both a and b are in S_1. Suppose a prefers b to y. Then the edge (a_-, b_+) is present in G_I since a prefers y (and thus b) to all neighbors in S_0; moreover, b_+ prefers a_- to $d(b)$. Hence (a_-, b_+) would be a blocking edge to N_I, contradicting its stability. So a prefers y to b. Similarly, b prefers z to a. Thus $\mathsf{wt}_M(a, b) = -2$.
2. Both a and b are in S_2. Either both a and b prefer their popular partners (y and z, resp.) to each other or the edge (a_0, b_0) is in G_I. In the latter case, (a_0, b_0) would be blocking edge to N_I since N_I contains $(a_0, d'(a))$ and $(b_0, d'(b))$. Thus both a and b prefer their popular partners to each other, so $\mathsf{wt}_M(a, b) = -2$.
3. One of a, b is in S_2 and the other is in S_1: assume wlog that $a \in S_2$ and $b \in S_1$. We claim that b prefers z to a. Otherwise the edge (a_+, b_-) would be in G_I since b prefers z (and thus a) to all neighbors in S_0. Note that (a_+, b_-) would block N_I since $(a_+, d(a)) \in N_I$. We next claim that a prefers y to b. Otherwise the edge (a_0, b_+) would be in G_I and this would be a blocking edge to N_I since $(a_0, d'(a))$ and $(b_+, d(b))$ are in N_I. Thus both a and b prefer their partners in M to each other, so $\mathsf{wt}_M(a, b) = -2$.

Hence in all cases, we have $\mathsf{wt}_M(a, b) = -2 = \alpha_a + \alpha_b$. ◁

▷ Claim 18. Any edge (a, b) where $\{\alpha_a, \alpha_b\} = \{-1, 0\}$ is covered.

Proof. Assume wlog $\alpha_a = -1$ and $\alpha_b = 0$. So $(a_-, y_+) \in N_I$ for some neighbor y of a. Also $(a_+, d(a)) \in N_I$. Observe that b_0 has to be matched in N_I, otherwise one of $(a_+, b_0), (a_0, b_0)$ – whichever is present in G_I – would be a blocking edge to N_I. So (z_0, b_0) is in N_I for some neighbor z of b.

If the edge (a_+, b_0) is present in E_I then it follows from the stability of N_I that b_0 prefers z_0 to a_+, i.e., b prefers z to a. Moreover, it follows from the existence of the edge (a_-, y_+) in E_I that a prefers y to all its neighbors in S_0, i.e., a prefers y to b if $b \in S_0$. If $b \in S_2$ and a prefers b to y then a prefers b to all neighbors in S_0 and so the edge (a_-, b_+) would have been present in E_I. This would have been a blocking edge to N_I since a_- prefers b_+ to y_+ and b_+ prefers a_- to $d(b)$. Thus a prefers y to b and so $\mathsf{wt}_M(a, b) = -2$.

The cases when (a_+, b_0) is *not* present in G_I are the following:

1. Both a, b are in S_2: there are two subcases here. In the first subcase, both a and b prefer their popular partners to each other and so $\mathsf{wt}_M(a, b) = -2$. In the second subcase, one of a, b prefers the other to its popular partner. Then the edge (a_0, b_0) is in E_I and the

stability of N_I implies that b_0 prefers z_0 to a_0 since $(a_0, d'(a)) \in N_I$. Thus b prefers z to a. This means that a prefers b to y and the edge (a_-, b_+) has to be in E_I since a prefers y (and thus b) to all neighbors in S_0. This makes (a_-, b_+) a blocking edge to N_I, a contradiction. Hence both a, b prefer their popular partners to each other, i.e., the second subcase does not arise. Thus $\mathsf{wt}_M(a, b) = -2$.

2. $b \in S_2$ prefers its popular partner to $a \in S_1$: so b prefers z to a and we have to argue that a prefers y to b. Suppose not, i.e., a prefers b to y. Since the edge (a_-, y_+) is in E_I, a prefers y (and thus b) to all neighbors in S_0. So the edge (a_-, b_+) is in E_I and this is a blocking edge to N_I since $(b_+, d(b))$ and (a_-, y_+) are in N_I. This contradicts N_I's stability, hence a prefers y to b. Thus $\mathsf{wt}_M(a, b) = -2$.
3. $a \in S_2$ prefers its popular partner to $b \in S_0$: so a prefers y to b. Then the edge (a_0, b_0) is in G_I. Since a_0 is matched to its least preferred neighbor $d'(a)$, it follows from the stability of N_I that b_0 prefers z_0 to a_0, i.e., b prefers z to a. Thus $\mathsf{wt}_M(a, b) = -2$.

Hence in all cases, we have $\mathsf{wt}_M(a, b) = -2 < -1 = \alpha_a + \alpha_b$. $\lhd$

This finishes the proof of Lemma 13. ◀

Finding a min-cost desired stable matching in G_I. We first check that all subscript $-$ vertices are stable in G_I. This is easily done by running Gale-Shapley algorithm in G_I and using the fact that all stable matchings leave the same vertices unmatched [15]. This ensures property (2). Then we solve a min-cost stable matching problem in G_I with *forbidden* edges. There are two types of forbidden edges here: the first type are all edges between a subscript $+$ vertex and a subscript 0 vertex in G_I. Forbidding these edges ensures property (1). The second type of forbidden edges are described below. Forbidding these edges ensures property (3).

Ensuring property (3). For any $u \in S_2$, all edges incident to any vertex u_+, u_0, u_- are marked forbidden *except* for the following edges, where v is u's popular partner:

- the edges among $(u_+, v_-), (u_0, v_0), (u_-, v_+)$ that are in E_I;
- the pair of edges $(u_+, d(u))$, $(u_0, d(u))$ and the pair of edges $(u_0, d'(u)), (u_-, d'(u))$.

For $u \in S_2$, every stable matching in G_I has to match $u_+, u_0, d(u), d'(u)$ since these are top choice neighbors for some vertices. Moreover, we have already checked that all subscript $-$ vertices are stable in G_I. Thus all the five vertices $u_+, u_0, u_-, d(u), d'(u)$ have to be matched in every stable matching in G_I. In particular, two of u_+, u_0, u_- are matched to $d(u), d'(u)$. Thus any stable matching in G_I that avoids forbidden edges of the second type has to contain one of $(u_+, v_-), (u_0, v_0), (u_-, v_+)$.

Desired stable matchings. We have seen that all stable matchings of G_I that satisfy the 3 properties given in Theorem 8 are precisely those stable matchings in G_I that avoid edges that we marked forbidden. Consider the stable matching polytope $\mathcal{S}$ of G_I: we know that $x_e \geq 0$ for any edge e is a valid inequality for $\mathcal{S}$, hence the intersection of $\mathcal{S}$ with the constraints $x_e = 0$ for every forbidden edge e is a face F of $\mathcal{S}$. Since F is an integral polytope and every integral point in F is a stable matching in G_I that avoids forbidden edges, N_I can be computed in polynomial time by linear programming over the constraints defining F. These are the constraints of the stable matching polytope $\mathcal{S}$ along with the constraints $x_e = 0$ for every forbidden edge e. A min-cost desired stable matching N_I over all $I \subseteq [p]$ maps to a min-cost popular matching in G (by Theorem 8 and Theorem 12).

As mentioned earlier, the popular subgraph G_0 can be constructed in linear time [7]. Then we identify the connected components $C_1, \ldots, C_p$ of size at least 4 in G_0. The number of sets I that we need to go through is 2^p, thus our algorithm runs in $2^p \cdot \mathsf{poly}(n)$ time. Since $p \le n/4$, this proves Theorem 2 stated in Section 1.

4 Semi-popular matchings

In this section we consider the problem of computing an *almost* semi-popular matching of cost at most $2\mathsf{opt}$. Our input is a bipartite graph $G = (A \cup B, E)$ where vertices have strict preferences and we have $\mathsf{cost} : E \to \mathbb{R}_{\ge 0}$. We are also given a parameter $\varepsilon \in (0, 1/2)$.

Popular fractional matchings. The notion of popularity can be extended to fractional matchings. A vector $\vec{x} \in \mathbb{R}^{|E|}_{\ge 0}$ that satisfies $\sum_{e \in \delta(u)} x_e \le 1$ for all vertices u is a fractional matching in G. The fractional matching $\vec{x}$ is popular if $\Delta(\vec{x}, N) \ge 0$ for all matchings N, where $\Delta(\vec{x}, N)$ is defined as follows: $\vec{x}$ is a convex combination of matchings (Birkhoff-von Neumann theorem), so $\vec{x} = \sum_i p_i I_{M_i}$ for some matchings M_i where $\sum_i p_i = 1$, each $p_i \ge 0$, and $\Delta(\vec{x}, N)$ is defined as $\sum_i p_i \cdot \Delta(M_i, N)$. Since the fractional matching $\vec{x}$ can possibly be expressed in multiple ways as convex combinations of matchings, $\Delta(\vec{x}, N)$ may seem ill-defined. However this is well-defined and we refer to [26, Lemma 1] for details.

Let opt^* be the cost of a min-cost popular fractional matching in G and let $\vec{q}$ be a min-cost popular fractional matching. The fractional matching $\vec{q}$ can be efficiently computed [26]. We have $\mathsf{cost}(\vec{q}) = \mathsf{opt}^* \le \mathsf{opt}$ where opt is the cost of a min-cost popular matching.

It was shown in [20] that the popular fractional matching polytope is half-integral. Thus we can assume that $\vec{q}$ is half-integral. So $\vec{q} = (I_{M_1} + I_{M_2})/2$ where M_1 and M_2 are two matchings in G. We know that $\Delta(\vec{q}, N) \ge 0$ for all matchings N in G.

▶ **Observation 3.** *There is a matching $M \in \{M_1, M_2\}$ such that M is semi-popular.*

Proof. Since $\Delta(\vec{q}, N) = (\Delta(M_1, N) + \Delta(M_2, N))/2$ and $\Delta(\vec{q}, N) \ge 0$ for every matching N, we have either $\Delta(M_1, N) \ge 0$ or $\Delta(M_2, N) \ge 0$ for every matching N. Hence one of M_1, M_2 is undefeated by at least half the matchings in G. ◁

Since all edge costs are non-negative and $\mathsf{cost}(\vec{q}) = (\mathsf{cost}(M_1) + \mathsf{cost}(M_2))/2$, we have $\mathsf{cost}(M_1) \le 2 \cdot \mathsf{cost}(\vec{q})$ and $\mathsf{cost}(M_2) \le 2 \cdot \mathsf{cost}(\vec{q})$. So there is $M \in \{M_1, M_2\}$ such that (i) M is semi-popular and (ii) $\mathsf{cost}(M) \le 2\mathsf{opt}$.

The problem here is to efficiently decide which of M_1, M_2 is semi-popular. We do not know how to answer this question exactly. However we can decide with high probability whether both M_1 and M_2 are close to being semi-popular or one of them is not - in which case the other matching has to be semi-popular (by Observation 3). Here we will use the classical result from [22] that shows a polynomial time algorithm to sample matchings from a distribution that is close to the uniform distribution in total variation distance (see [22, Corollary 4.3]).

The input is $G = (A \cup B, E)$ with non-negative edge costs and $\varepsilon \in (0, 1/2)$. Our algorithm is as follows:

1. Compute a min-cost popular half-integral matching $\vec{q}$ in G. Let $\vec{q} = (I_{M_1} + I_{M_2})/2$ where M_1 and M_2 are matchings in G.
2. Produce a sample $\mathcal{S}$ of $s = 64 \cdot \lceil (\ln n)/\varepsilon^2 \rceil$ matchings from a distribution that is $\varepsilon/4$-close to the uniform distribution (on all matchings in G) in total variation distance.
3. If both M_1 and M_2 are undefeated by more than $s \cdot (1 - \varepsilon)/2$ of matchings in $\mathcal{S}$ then return the matching in $\{M_1, M_2\}$ with lower cost.
4. Else return the matching in $\{M_1, M_2\}$ undefeated by a majority of matchings in $\mathcal{S}$.

In Step 2, we use the random sampler in [22] that constructs the sample $\mathcal{S}$ in $\mathsf{poly}(n, \frac{1}{\varepsilon})$ time. It is easy to see that the running time of our algorithm is $\mathsf{poly}(n, \frac{1}{\varepsilon})$. Lemma 19 and Lemma 20 bound the probability that our algorithm makes an error.

▶ **Lemma 19.** *Suppose $M \in \{M_1, M_2\}$ is defeated by more than $1/2 + \varepsilon$ fraction of all matchings in G. Then our algorithm returns M in step 3 with probability at most $1/n$.*

Proof. Since M is defeated by more than $1/2+\varepsilon$ fraction of all matchings in G, the expected number of matchings that defeat M from a set of s matchings, where each matching is chosen uniformly at random from the set of all matchings in G is more than $s \cdot (1/2 + \varepsilon)$. The set $\mathcal{S}$ is formed by sampling s matchings from a distribution $\varepsilon/4$-close to the uniform distribution in total variation distance. Hence the expected number of matchings from $\mathcal{S}$ that defeat M is more than $s \cdot (1/2 + \varepsilon - \varepsilon/4) = s \cdot (2 + 3\varepsilon)/4$.

If M was returned in step 3 then M was undefeated by more than $s \cdot (1 - \varepsilon)/2$ matchings from $\mathcal{S}$. Equivalently, less than $s \cdot (1 + \varepsilon)/2$ matchings from $\mathcal{S}$ defeated M. By Chernoff bound, the probability of this event is at most $\exp(-s \cdot \varepsilon^2/(16+24\varepsilon))$. Since $s \geq 64 \cdot (\ln n)/\varepsilon^2$, this probability is at most $1/n$. ◀

The next lemma bounds the error when our algorithm reaches step 4.

▶ **Lemma 20.** *Suppose $M \in \{M_1, M_2\}$ is not semi-popular. Then our algorithm returns M in step 4 with probability at most $1/n$.*

Proof. Since M is defeated by more than half the matchings in G, the expected number of matchings that defeat M from a set of s matchings, where each matching is chosen uniformly at random from the set of all matchings in G, is more than $s/2$. Since the set $\mathcal{S}$ is formed by sampling s matchings from a distribution $\varepsilon/4$-close to the uniform distribution in total variation distance, the expected number of matchings that defeat N from $\mathcal{S}$ is more than $s \cdot (2 - \varepsilon)/4$.

The algorithm reached step 4 and M was the matching that was undefeated by a majority of matchings in $\mathcal{S}$. Observe that M defeated more than $s \cdot (1 + \varepsilon)/2$ matchings in the set $\mathcal{S}$. This is because the matching in $\{M_1, M_2\} \setminus \{M\}$ was defeated by more than $s \cdot (1 + \varepsilon)/2$ matchings in $\mathcal{S}$ – otherwise we would not have reached step 4. Since M defeats more than $s(1+\varepsilon)/2$ matchings from $\mathcal{S}$, less than $s(1-\varepsilon)/2$ matchings from $\mathcal{S}$ defeated M. By Chernoff bound, the probability of this event is at most $\exp(-s \cdot \varepsilon^2/(64-32\varepsilon))$. Since $s \geq 64 \cdot (\ln n)/\varepsilon^2$, this probability is at most $1/n$. ◀

Lemma 19 and Lemma 20 bound the error probability of our algorithm. Thus we have proved Theorem 4 stated in Section 1.

References

1 D. J. Abraham, R. W. Irving, T. Kavitha, and K. Mehlhorn. Popular matchings. *SIAM Journal on Computing*, 37(4):1030–1045, 2007.

2 P. Biro, R. W. Irving, and D. F. Manlove. Popular matchings in the marriage and roommates problems. In *Proceedings of the seventh International Conference on Algorithms and Complexity (CIAC)*, pages 97–108, 2010.

3 M.-J.-A.-N. de C. (Marquis de) Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* L'Imprimerie Royale, 1785.

4 Condorcet method. `https://en.wikipedia.org/wiki/Condorcet_method`.

5 Á. Cseh. Popular matchings. Trends in Computational Social Choice, Ulle Endriss (ed.), 2017.

6 Á. Cseh, C.-C. Huang, and T. Kavitha. Popular matchings with two-sided preferences and one-sided ties. *SIAM Journal on Discrete Mathematics*, 31(4):2348–2377, 2017.

7 Á. Cseh and T. Kavitha. Popular edges and dominant matchings. *Mathematical Programming*, 172(1):209–229, 2018.

8 Y. Faenza and T. Kavitha. Quasi-popular matchings, optimality, and extended formulations. In *Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 325–344, 2020.

9 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.

10 T. Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45(2):233–284, 1992.

11 T. Feder. Network flow and 2-satisfiability. *Algorithmica*, 11(3):291–319, 1994.

12 T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28(1):103–126, 2003.

13 F. V. Fomin and D. Kratsch. *Exact exponential algorithms*. Springer-Verlag New York, Inc., New York, 2010.

14 D. Gale and L.S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

15 D. Gale and M. Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11:223–232, 1985.

16 P. Gärdenfors. Match making: assignments based on bilateral preferences. *Behavioural Science*, 20:166–173, 1975.

17 S. Gupta, P. Misra, S. Saurabh, and M. Zehavi. Popular matching in roommates setting is np-hard. In *Proceedings of the 30th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2810–2822, 2019.

18 M. Hirakawa, Y. Yamauchi, S. Kijima, and M. Yamashita. On the structure of popular matchings in the stable marriage problem - who can join a popular matching? In the 3rd International Workshop on Matching Under Preferences (MATCH-UP), 2015.

19 C.-C. Huang and T. Kavitha. Popular matchings in the stable marriage problem. *Information and Computation*, 222:180–194, 2013.

20 C.-C. Huang and T. Kavitha. Popularity, mixed matchings, and self-duality. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2294–2310, 2017.

21 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.

22 M. Jerrum and A. Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989.

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

24 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.

25 T. Kavitha. Popular roommates in simply exponential time. In *Proceedings of the 39th Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 20:1–20:15, 2019.

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

27 K. Makarychev, Y. Makarychev, M. Sviridenko, and J. Ward. A bi-criteria approximation algorithm for k-means. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM)*, pages 14:1–14:20, 2016.

28 E. McDermid and R. W. Irving. Sex-equal stable matchings: Complexity and exact algorithms. *Algorithmica*, 68(3):545–570, 2014.

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

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

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