Popularity in the generalized Hospital Residents setting

Meghana Nasre and Amit Rawat

Indian Institute of Technology Madras, India

Abstract. We consider the problem of computing *popular* matchings in a bipartite graph $G = (\mathcal{R} \cup \mathcal{H}, E)$ where $\mathcal{R}$ and $\mathcal{H}$ denote a set of residents and a set of hospitals respectively. Each hospital h has a positive capacity denoting the number of residents that can be matched to h. The residents and the hospitals specify strict preferences over each other. This is the well-studied Hospital Residents (HR) problem which is a generalization of the Stable Marriage (SM) problem. The goal is to assign residents to hospitals *optimally* while respecting the capacities of the hospitals. Stability is a well-accepted notion of optimality in such problems. However, motivated by the need for larger cardinality matchings, alternative notions of optimality like *popularity* have been investigated in the SM setting. In this paper, we consider a generalized HR setting – namely the Laminar Classified Stable Matchings (LCSM$^+$) problem. Here, additionally, hospitals can specify classifications over residents in their preference lists and classes have upper quotas. We show the following new results: We define a notion of popularity and give a structural characterization of popular matchings for the LCSM$^+$ problem. Assume $n = |\mathcal{R}|+|\mathcal{H}|$ and $m = |E|$. We give an $O(mn)$ time algorithm for computing a maximum cardinality popular matching in an LCSM$^+$ instance. We give an $O(mn^2)$ time algorithm for computing a matching that is popular amongst the maximum cardinality matchings in an LCSM$^+$ instance.

1 Introduction

Consider an academic institution where students credit an elective course from a set of available courses. Every student and every course rank a subset of elements from the other set in a strict order of preference. Each course has a quota denoting the maximum number of students it can accommodate. The goal is to allocate to every student at most one course respecting the preferences. This is the well-studied Hospital Residents problem [7]. We consider its generalization where, in addition, a course can *classify* students − for example, the students may be classified as under-graduates and post-graduates and department-wise and so on. Depending on the classifications, a student may belong to multiple classes. Apart from the total quota, each course now has a quota for every class. An allocation, in this setting, has to additionally respect the class quotas. This is the Classified Stable Matching problem introduced by Huang [10].

Stability is a de-facto notion of optimality in settings where both set of participants have preferences. Informally, an allocation of students to courses is stable if no unallocated student-course pair has incentive to deviate from the allocation. Stability is appealing for several reasons − stable allocations are guaranteed to exist, they are efficiently computable and all stable allocations leave the same set of students unallocated [9]. However, it is known [13] that the cardinality of a stable allocation can be half the size of the largest sized allocation possible. Furthermore, in applications like student-course allocation, leaving a large number of students unallocated is undesirable. Thus, it is interesting to consider notions of optimality which respect preferences but possibly compromise stability in the favor of cardinality. Kavitha and Huang [11,13] investigated this in the Stable Marriage (SM) setting where they considered *popularity* as an alternative to stability. At a high level, an allocation of students to courses is *popular* if no *majority* wishes to deviate from the allocation. Here, we consider popularity in the context of two-sided preferences and one-sided capacities with classifications.

We formally define our problem now – we use the familiar hospital residents notation. Let $G = (\mathcal{R} \cup \mathcal{H}, E)$ be a bipartite graph where $|\mathcal{R} \cup \mathcal{H}| = n$ and $|E| = m$. Here $\mathcal{R}$ denotes the set of residents, $\mathcal{H}$ denotes the set of hospitals and every hospital $h \in \mathcal{H}$ has an upper quota $q(h)$ denoting the maximum number of residents h can occupy. A pair $(r, h) \in E$ denotes that r and h are mutually acceptable to each other. Each resident (resp. hospital) has a strict ordering of a subset of the hospitals (resp. residents) that are acceptable to him or her (resp. it). This ordering is called the preference list of a vertex. An assignment (or a matching) M in G is a subset of E such that every resident is assigned to at most one hospital and a hospital h is assigned at most $q(h)$ residents. Let $M(r)$ (resp. $M(h)$) denote the hospital (resp. the set of residents) which are assigned to r (resp. h) in M. A hospital h is under-subscribed if $|M(h)| < q(h)$. A matching M is *stable* if no unassigned pair (r, h) wishes to deviate from M. The goal is to compute a stable matching in G. We denote it by HR$^+$ throughout the paper [1]. The celebrated deferred acceptance algorithm by Gale and Shapley [7] proves that every instance of the HR$^+$ problem admits a stable matching.

A generalization of the HR$^+$ problem is the Laminar Classified Stable Matching (LCSM) problem introduced by Huang [10]. An instance of the LCSM$^+$ problem is an instance of the HR$^+$ problem where additionally, each hospital h is allowed to specify a classification over the set of residents in its preference list. A class C^h_k of a hospital h is a subset of residents in its preference list and has an associated upper quota $q(C^h_k)$ denoting the maximum number of residents that can be matched to h in C^h_k. (In the LCSM problem [10], classes can have lower quotas as well.) We assume that the classes of a hospital form a *laminar* set. That is, for any two classes C^h_j and C^h_k, either the two classes are disjoint ($C^h_j \cap C^h_k = \emptyset$), or one is contained inside the other ($C^h_j \subset C^h_k$ or $C^h_k \subset C^h_j$). Huang suitably modified the classical definition of stability to account for the presence of these classifications. He showed that every instance of the LCSM$^+$ problem admits a stable matching which can be computed in $O(mn)$ time [10]. A restriction of the LCSM$^+$ problem, denoted by Partition Classified Stable Matching (PCSM$^+$), is where the classes of every hospital partition the residents in its preference list.

Motivated by the need to output larger cardinality matchings, we consider computing *popular* matchings in the LCSM$^+$ problem. The notion of popularity uses *votes* to compare two matchings. Before we can define voting in the LCSM$^+$ setting, it is useful to discuss voting in the context of the SM problem.

Voting in the SM setting: Let $G = (\mathcal{R} \cup \mathcal{H}, E)$ be an instance of the SM problem and let M and M' be any two matchings in G. A vertex $u \in \mathcal{R} \cup \mathcal{H}$ (where each hospital h has $q(h) = 1$) prefers M over M' and therefore votes for M over M' if either (i) u is matched in M and unmatched in M' or (ii) u is matched in both M and M' and prefers $M(u)$ over $M'(u)$. A matching M is more popular than M' if the number of votes that M gets as compared to M' is greater than the number of votes that M' gets as compared to M. A matching M is popular if there does not exist any matching that is more popular than M. In the SM setting it is known that a stable matching is popular, however it was shown to be *minimum* cardinality popular matching [11]. Huang and Kavitha [11,13] gave efficient algorithms for computing a max-cardinality popular matching and a popular matching amongst max-cardinality matchings in an SM instance.

Voting in the capacitated setting: To extend voting in the capacitated setting, we assign a hospital h as many votes as its upper quota $q(h)$. This models the scenario in which hospitals with larger capacity get a larger share of votes. For the HR$^+$ problem, a hospital h compares the most preferred resident in $M(h) \setminus M'(h)$ to the most preferred resident in $M'(h) \setminus M(h)$ (and votes for M or M' as far as those two residents are concerned) and so on. For this voting scheme, we can obtain analogous results for computing popular matchings in the HR$^+$ problem via the standard technique of *cloning* (that is, creating $q(h)$ copies of a hospital h and appropriately modifying preference lists of the residents and hospitals [2]). However, our interest is in the LCSM$^+$ problem, for which we are not aware of any reduction to the SM problem. Furthermore, we show that the straightforward voting scheme as defined in the HR$^+$ does not suffice for the LCSM$^+$ problem. Therefore, we define a voting scheme for a hospital which takes into consideration the classifications as well as ensures that every stable matching in the LCSM$^+$ instance is popular. We show the following results:

- We define a notion of popularity for the LCSM$^+$ problem. Since our definition ensures that stable matchings are popular – this guarantees the existence of popular matchings in the LCSM$^+$ problem.
- We give a characterization of popular matchings for the LCSM$^+$ problem, which is a natural extension of the characterization of popular matchings in SM setting [11].

[1] We use HR$^+$ instead of HR for consistency with other problems discussed in the paper.

[2] For every hospital in the cloned graph, its preference list is the same as in the original instance. For every hospital h, fix an ordering of its clones. The preference list of a resident r in the cloned instance is obtained by replacing the occurrence of h by the fixed ordering of its clones. We refer the reader to [4,14] for details.

– We obtain the following algorithmic results. An $O(m+n)$ (resp. $O(mn)$) time algorithm for computing a maximum cardinality popular matching in a PCSM$^+$ (resp. LCSM$^+$) instance. An $O(mn)$ (resp. $O(mn^2)$) time algorithm for computing a popular matching amongst maximum cardinality matchings in a PCSM$^+$ (resp. LCSM$^+$) instance.

Very recently, independent of our work, two different groups [4,12] have considered popular matchings in the one-to-many setting. Brandl and Kavitha [4] have considered computing *popular* matchings in the HR$^+$ problem. In their work as well as ours, a hospital h is assigned as many votes as its capacity to compare two matchings M and M'. In contrast, by the definition of popularity in [4], a hospital h chooses the most adversarial ordering of residents in $M(h) \setminus M'(h)$ and $M'(h) \setminus M(h)$ for comparing M and M'. However, it is interesting to note that in an HR$^+$ instance the same matching is output by both our algorithms. On the other hand, we remark that the model considered in our paper is a more general one than the one considered in [4]. Kamiyama [12] has generalized our work and the results in [4] using a matroid based approach.

We finally remark that one can consider voting schemes where a hospital is given a *single* vote instead of capacity many votes. In one such scheme, a hospital compares the set of residents in $M(h)$ and $M'(h)$ in lexicographic order and votes accordingly. However, when such a voting is used, it is possible to construct instances where a stable matching is *not popular*. The techniques in this paper use the fact that stable matchings are popular, therefore it is unclear if our techniques will apply for such voting schemes.

Related Work: The notion of popularity was introduced by Gärdenfors [8] in the context of stable matchings. In [1] Abraham et al. studied popularity in the one-sided preference list model. As mentioned earlier, our work is inspired by a series of papers where popularity is considered as an alternative to stability in the stable marriage setting by Huang, Kavitha and Cseh [5,11,13]. Biró et al. [3] give several practical scenarios where stability may be compromised in the favor of size. The PCSM$^+$ problem is a special case of the Student Project Allocation (SPA) problem studied by Abraham et al. [2]. They gave a linear time algorithm to compute a stable matching in an instance of the SPA problem. In this paper, we use the algorithms of Abraham et al. [2] and Huang [10] for computing stable matchings in the PCSM$^+$ and LCSM$^+$ problems. Both these algorithms follow the standard *deferred acceptance* algorithm of Gale and Shapley with problem specific modifications. We refer the reader to [2] and [10] for details.

Organization: In Section 2 we define the notion of popularity, in Section 3 we present the structural characterization of popular matchings. In Section 4 we describe our algorithms to compute a maximum cardinality popular matching, and a popular matching amongst maximum cardinality matchings. We conclude with a short discussion about popular matchings in the LCSM problem.

2 Stability and popularity in the LCSM$^+$ problem

Consider an instance $G = (\mathcal{R} \cup \mathcal{H}, E)$ of the LCSM$^+$ problem. As done in [10], assume that for every $h \in \mathcal{H}$ there is a class C_*^h containing all the residents in the preference list of h and $q(C_*^h) = q(h)$. For a hospital h, let $T(h)$ denote the tree of classes corresponding to h where C_*^h is the root of $T(h)$. The leaf classes in $T(h)$ denote the most refined classifications for a resident whereas as we move up in the tree from a leaf node to the root, the classifications gets coarser.

To define stable matchings in the LCSM problem, Huang introduced the notion of a *blocking group* w.r.t. a matching. Later, Fleiner and Kamiyama [6] defined a blocking pair which is equivalent to a blocking group of Huang. We use the definition of stability from [6] which we recall below. A set $S = \{r_1, \ldots, r_l\}$ is *feasible* for a hospital h if $|S| \le q(h)$ and for every class C_j^h of h (including the root class C_*^h), we have $|C_j^h \cap S| \le q(C_j^h)$. A matching M in G is feasible if every resident is matched to at most one hospital, and $M(h)$ is feasible for every hospital $h \in \mathcal{H}$. A pair $(r, h) \notin M$ blocks M iff both the conditions below hold:

– r is unmatched in M, or r prefers h over $M(r)$, and
– either the set $M(h) \cup \{r\}$ is feasible for h, or there exists a resident $r' \in M(h)$, such that h prefers r over r', and $(M(h) \setminus \{r'\}) \cup \{r\}$ is feasible for h.

A feasible matching M in G is stable if M does not admit any blocking pair.

2.1 Popularity

To define popularity, we need to specify how a hospital compares two sets $M(h)$ and $M'(h)$ in an LCSM$^+$ setting, where M and M' are two feasible matchings in the instance.

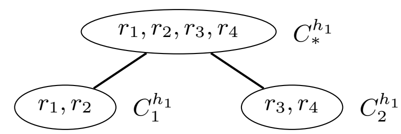
Illustrative example Consider the following LCSM$^+$ instance where $\mathcal{R} = \{r_1, \ldots, r_4\}$ and $\mathcal{H} = \{h_1, \ldots, h_3\}$ and the preference lists of the residents and hospitals are as given in Figure 1(a) and (b) respectively. The preferences can be read as follows: resident r_1 has h_1 as his top choice hospital. Resident r_2 has h_2 as its top choice hospital followed by h_1 which is his second choice hospital and so on. For $h \in \{h_2, h_3\}$ we have $q(h) = 1$ and both these hospitals have a single class C_*^h containing all the residents in the preference list of h and $q(C_*^h) = q(h)$. For hospital h_1 we have $q(h_1) = 2$ and the classes provided by h_1 are $C_1^{h_1} = \{r_1, r_2\}, C_2^{h_1} = \{r_3, r_4\}, C_*^{h_1} = \{r_1, r_2, r_3, r_4\}$ with quotas as follows: $q(C_1^{h_1}) = q(C_2^{h_1}) = 1$ and $q(C_*^{h_1}) = 2$. We remark that the example in Figure 1 is also a PCSM$^+$ instance. Figure 1(c) shows the tree $T(h_1)$.

$\mathbf{r_1}$: h_1
$\mathbf{r_2}$: h_2, h_1, h_3
$\mathbf{r_3}$: h_1, h_2
$\mathbf{r_4}$: h_1

(a)

$\mathbf{h_1}$: r_2, r_3, r_4, r_1
$\mathbf{h_2}$: r_3, r_2
$\mathbf{h_3}$: r_2

(b)



(c)

Fig. 1. (a) Resident preferences, (b) Hospital preferences, (c) $T(h_1)$. The matchings $M = \{(r_1, h_1), (r_2, h_2), (r_3, h_1)\}$, $M' = \{(r_2, h_1), (r_3, h_2), (r_4, h_1)\}$, and $M'' = \{(r_1, h_1), (r_2, h_3), (r_3, h_2), (r_4, h_1)\}$ are all feasible in the instance.

Consider the two feasible matchings M and M' defined in Fig. 1. Note that M is stable in the instance whereas the edge (r_3, h_1) blocks M'. While comparing M and M', the vote for every vertex u in the instance except h_1 is clear – u compares $M(u)$ with $M'(u)$ and votes accordingly. In order for h_1 to vote between M and M', the hospital compares between $M(h_1) = \{r_1, r_3\}$ and $M'(h_1) = \{r_2, r_4\}$. A straightforward way is to compare r_3 with r_2 (the most preferred resident in $M(h_1)$ to the most preferred resident in $M'(h_1)$) and then compare r_1 with r_4 (second most preferred resident in $M(h_1)$ to second most preferred resident in $M'(h_1)$). Thus, both the votes of h_1 are in favor of M' when compared with M. Such a comparison has two issues – (i) it ignores the classifications given by h_1, and (ii) the number of votes that M' gets when compared with M is more than the number of votes that M gets as compared to M'. Therefore M' is more popular than M which implies that M (a stable matching) is **not** popular.

We propose a comparison scheme for hospitals which addresses both the issues. In the above example, we note that $r_1 \in M(h)$ has a corresponding resident $r_2 \in M'(h)$ to be compared to in one of the most refined classes $C_1^{h_1}$ (see Figure 1(c)). Thus, we compare r_1 with r_2. The resident $r_3 \in M(h)$ is compared to $r_4 \in M(h)$ another leaf class $C_2^{h_1}$. According to this comparison, h_1 is indifferent between M and M' and M' is no longer more popular than M. Note that, although in the example, both the comparisons happen in a leaf class, this may not be the case in a general instance. Finally, we note that the matching M'' is a popular matching in the instance and is strictly larger in size than the stable matching M.

We formalize the above observations in the rest of the section. To take into account the classifications, for a hospital h and the matchings M and M', we set up a correspondence between residents in $M(h) \setminus M'(h)$ and the residents in $M'(h) \setminus M(h)$. That is, we define:

$$\mathbf{corr} : M(h) \oplus M'(h) \to M(h) \oplus M'(h) \cup \{\perp\}$$

For a resident $r \in M(h) \oplus M'(h)$ we denote by $\mathbf{corr}(r)$ the corresponding resident to which r gets compared when the hospital h casts its votes. We let $\mathbf{corr}(r) = \perp$ if r does not have a corresponding resident to be compared to from the other matching. The pseudo-code for the algorithm to compute the **corr** function is given below.

Algorithm 1 Correspondence between residents of $M(h)$ and $M'(h)$

procedure FIND-CORRESPONDENCE(h, M, M')
  let $T(h)$ be the classification tree associated with h
  set $\mathbf{corr}(r) = \perp$ for each $r \in M(h) \oplus M'(h)$
  $Y = M(h) \setminus M'(h)$; $Y' = M'(h) \setminus M(h)$
  while $Y \neq \emptyset$ and $Y' \neq \emptyset$ **do**
    for each class C_j^h in $T(h)$ **do**
      $X_j = C_j^h \cap Y$
      $X'_j = C_j^h \cap Y'$
    Let C_f^h be one of the most refined classes for which $X_f \neq \emptyset$ and $X'_f \neq \emptyset$.
    for $k = 1, \ldots, \min(|X_f|, |X'_f|)$ **do**
      let r be the k-th most preferred resident in X_f
      let r' be the k-th most preferred resident in X'_f
      set $\mathbf{corr}(r) = r'$, and $\mathbf{corr}(r') = r$
      $Y = Y \setminus \{r\}$; $Y' = Y' \setminus \{r'\}$

The algorithm begins by setting **corr** for every $r \in M(h) \oplus M'(h)$ to $\perp$. The algorithm maintains two sets of residents $Y = M(h) \setminus M'(h)$ and $Y' = M'(h) \setminus M(h)$ for whom **corr** needs to be set. As long as the sets Y and Y' are both non-empty, the algorithm repeatedly computes for every class C_j^h (including the root class C_*^h) the sets $X_j = C_j^h \cap Y$ and $X'_j = C_j^h \cap Y'$. The algorithm then chooses one of the most refined classes, say C_f^h in $T(h)$, for whom X_f and X'_f are both non-empty. Finally, residents in X_f and X'_f are sorted according to the preference ordering of h and the **corr** of the k-th most preferred resident in X_f is set to the k-th most preferred resident in X'_f, where $k = 1, \ldots, \min\{|X_f|, |X'_f|\}$.

For $r \in \mathcal{R}$, and any feasible matching M in G, if r is unmatched in M then, $M(r) = \perp$. A vertex prefers any of its neighbours over $\perp$. For a vertex $u \in \mathcal{R} \cup \mathcal{H}$, let $x, y \in N(u) \cup \{\perp\}$, where $N(u)$ denotes the neighbours of u in G.

$$\begin{aligned} vote_u(x, y) &= +1 \quad \text{if } u \text{ prefers } x \text{ over } y \\ &= -1 \quad \text{if } u \text{ prefers } y \text{ over } x \\ &= \;\;\, 0 \quad \text{if } x = y \end{aligned}$$

Using the above notation, the vote of a resident is easy to define – a resident r prefers M' over M iff the term $\mathcal{V}_r > 0$, where $\mathcal{V}_r = vote_r(M'(r), M(r))$.

Recall that a hospital h uses $q(h)$ votes to compare M and M'. Let $q_1(h) = |M(h) \cap M'(h)|$ (number of common residents assigned to h in M and M') and $q_2(h) = q(h) - \max\{|M(h)|, |M'(h)|\}$ (number of unfilled positions of h in both M and M'). Our voting scheme ensures that $q_1(h) + q_2(h)$ votes of h remain unused when comparing M and M'. A hospital h prefers M' over M iff the term $\mathcal{V}_h > 0$, where $\mathcal{V}_h$ is defined as follows:

$$\mathcal{V}_h = (|M'(h)| - |M(h)|) + \sum_{\substack{r \in M'(h) \setminus M(h) \\ \&\& \\ \mathbf{corr}(r) \neq \perp}} vote_h(r, \mathbf{corr}(r))$$

The first term in the definition of $\mathcal{V}_h$ counts the votes of h w.r.t. the residents from either M or M' that did not find correspondence. The second term counts the votes of h w.r.t. the residents each of which has a corresponding resident from the other matching. We note that in the SM setting, $\mathbf{corr}(r)$ will simply be $M(h)$. Thus, our definition of votes in the presence of capacities is a natural generalization of the voting scheme in the SM problem.

Let us define the term $\Delta(M', M)$ as the difference between the votes that M' gets over M and the votes that M gets over M'.

$$\Delta(M', M) = \sum_{r \in \mathcal{R}} \mathcal{V}_r + \sum_{h \in \mathcal{H}} \mathcal{V}_h$$

Definition 1. *A matching M is popular in G iff for every feasible matching M', we have $\Delta(M', M) \leq 0$.*

2.2 Decomposing $M \oplus M'$

Here, we present a simple algorithm which allows us to decompose edges of components of $M \oplus M'$ in an instance into alternating paths and cycles. Consider the graph $\tilde{G} = (\mathcal{R} \cup \mathcal{H}, M \oplus M')$, for any two feasible matchings M and M' in G. We note that the degree of every resident in $\tilde{G}$ is at most 2 and the degree of every hospital in $\tilde{G}$ is at most $2 \cdot q(h)$. Consider any connected component $\mathcal{C}$ of $\tilde{G}$ and let e be any edge in $\mathcal{C}$. We observe that it is possible to construct a unique maximal M alternating path or cycle ρ containing e using the following simple procedure. Initially ρ contains only the edge e.

1. Let $r \in \mathcal{R}$ be one of the end points of the path ρ, and assume that $(r, M(r)) \in \rho$. We grow ρ by adding the edge $(r, M'(r))$. Similarly if an edge from M' is incident on r in ρ, we grow the path by adding the edge $(r, M(r))$ if it exists.
2. Let $h \in \mathcal{H}$ be one of the end points of the path ρ, and assume that $(r, h) \in M \setminus M'$ belongs to ρ. We extend ρ by adding $(\mathbf{corr}(r), h)$ if $\mathbf{corr}(r)$ is not equal to $\perp$. A similar step is performed if the last edge on ρ is $(r, h) \in M' \setminus M$.
3. We stop the procedure when we complete a cycle (ensuring that the two adjacent residents of a hospital are **corr** for each other according to the hospital), or the path can no longer be extended. Otherwise we go to Step 1 or Step 2 as applicable and repeat.

The above procedure gives us a unique decomposition of a connected component in $\tilde{G}$ into alternating paths and cycles. Note that a hospital may appear multiple times in a single path or a cycle and also can belong to more than one alternating paths and cycles. Figure 2 gives an example of the decomposition of the two feasible matchings in the instance in Figure 1.

Fig. 2. M and M' are feasible matchings in the example as defined in Fig. 1. (a) $\tilde{G} = (\mathcal{R} \cup \mathcal{H}, M \oplus M')$; bold edges belong to M, dashed edges belong to M'. (b) shows the decomposition of the edges of the component of $\tilde{G}$ into a single path.

Let $\mathcal{Y}_{M \oplus M'}$ denote the collection of alternating paths and alternating cycles obtained by decomposing every component of $\tilde{G}$.

We now state a useful property about any alternating path or cycle in $\mathcal{Y}_{M \oplus M'}$.

Lemma 1. *If ρ is an alternating path or an alternating cycle in $\mathcal{Y}_{M \oplus M'}$, then $M \oplus \rho$ is a feasible matching in G.*

Proof. Let $\langle r', h, r \rangle$ be any sub-path of ρ, where $r' = \mathbf{corr}(r)$, and $(r, h) \in M$. We prove that $(M(h) \setminus \{r\}) \cup \{r'\}$ is feasible for h. Let C_i^h (resp. C_j^h) be the unique leaf class of $T(h)$ containing r (resp. r'). See Figure 3. We consider the following two cases:

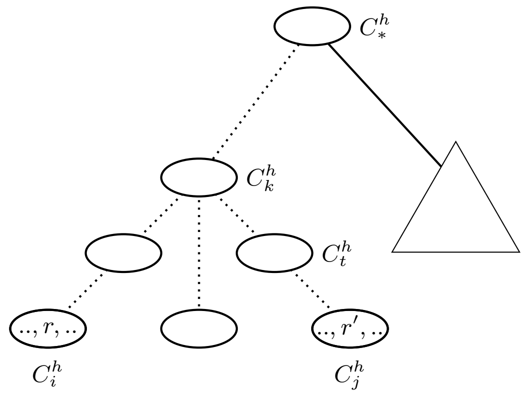


Fig. 3. The classification tree T(h) for a hospital h.

- r and r' belong to the same leaf class in $T(h)$, i.e. $C_i^h = C_j^h$. In this case, it is easy to note that $(M(h) \setminus \{r\}) \cup \{r'\}$ is feasible for h.
- r and r' belong to different leaf classes of $T(h)$, i.e. $C_i^h \neq C_j^h$. Observe that $|(M(h) \setminus \{r\}) \cup \{r'\}|$ can violate the upper quota only for those classes of $T(h)$ which contain r' but do not contain r. Let C_k^h be the least common ancestor of C_i^h and C_j^h in $T(h)$. It suffices to look at any class C_t^h which lies in the path from C_k^h to C_j^h excluding the class C_k^h and show that $|(M(h) \cap C_t^h) \cup \{r'\}| \leq q(C_t^h)$. As $r' = \mathbf{corr}(r)$ and $r \notin C_t^h$, we claim that $|M(h) \cap C_t^h| < |M'(h) \cap C_t^h| \leq q(C_t^h)$. The first inequality is due to the fact that r' did not find a corresponding resident in the set $(M(h) \setminus M'(h)) \cap C_t^h$. The second inequality is because M' is feasible. Thus, $(M(h) \cap C_t^h) \cup \{r'\}$ does not violate the upper quota for C_t^h. Therefore $(M(h) \setminus \{r\}) \cup \{r'\}$ is feasible for h.

We note that the hospital h may occur multiple times on ρ. Let $M(h)_\rho$ denote the set of residents matched to h restricted to ρ. To complete the proof of the Lemma, we need to prove that $(M(h) \setminus M(h)_\rho) \cup M'(h)_\rho$ is feasible for h. The arguments for this follow from the arguments given above. □

As was done in [13], it is convenient to label the edges of $M' \setminus M$ and use these labels to compute $\Delta(M', M)$. Let $(r, h) \in M' \setminus M$; the label on (r, h) is a tuple:

$$(vote_r(h, M(r)), \quad vote_h(r, \mathbf{corr}(r)))$$

Note that since we are labeling edges of $M' \setminus M$, both entries of the tuple come from the set $\{-1, 1\}$. With these definitions in place, we are ready to give the structural characterization of popular matchings in an LCSM$^+$ instance.

3 Structural characterization of popular matchings

Let $G = (\mathcal{R} \cup \mathcal{H}, E)$ be an LCSM$^+$ instance and let M and M' be two feasible matchings in G. Using the **corr** function, we obtain a correspondence of residents in $M(h) \oplus M'(h)$ for every hospital h in G. Let $\tilde{G} = (\mathcal{R} \cup \mathcal{H}, M \oplus M')$ and let $\mathcal{Y}_{M \oplus M'}$ denote the collection of alternating paths and cycles obtained by decomposing every component of $\tilde{G}$. Finally, we label the edges of $M' \setminus M$ using appropriate votes. The goal of these steps is to is to rewrite the term $\Delta(M', M)$ as a sum of labels on edges.

We note that the only vertices for whom their vote does not get captured on the edges of $M' \setminus M$ are vertices that are matched in M but not matched in M'. Let $\mathcal{U}$ denote the multi-set of vertices that are end points of paths in $\mathcal{Y}_{M \oplus M'}$ such that there is no M' edge incident on them. Note that the same hospital can belong to multiple alternating paths and cycles in $\mathcal{Y}_{M \oplus M'}$, therefore we need a multi-set. All vertices in $\mathcal{U}$

prefer M over M' and hence we add a -1 while capturing their vote in $\Delta(M', M)$. We can write $\Delta(M', M)$ as:

$$\Delta(M', M) = \sum_{x \in \mathcal{U}} -1 + \sum_{\rho \in \mathcal{Y}_{M \oplus M'}} \left(\sum_{(r,h) \in (M' \cap \rho)} \{vote_r(h, M(r)) + vote_h(r, \mathbf{corr}(r))\} \right)$$

We now delete the edges labeled $(-1, -1)$ from all paths and cycles ρ in $\mathcal{Y}_{M \oplus M'}$. This simply breaks paths and cycles into one or more paths. Let this new collection of paths and cycles be denoted by $\tilde{\mathcal{Y}}_{M \oplus M'}$. Let $\tilde{\mathcal{U}}$ denote the multi-set of vertices that are end points of paths in $\tilde{\mathcal{Y}}_{M \oplus M'}$ such that there is no M' edge incident on them. We rewrite $\Delta(M', M)$ as:

$$\Delta(M', M) = \sum_{x \in \tilde{\mathcal{U}}} -1 + \sum_{\rho \in \tilde{\mathcal{Y}}_{M \oplus M'}} \left(\sum_{(r,h) \in (M' \cap \rho)} \{vote_r(h, M(r)) + vote_h(r, \mathbf{corr}(r))\} \right)$$

Theorem below characterizes a popular matching.

Theorem 1. *A feasible matching M in G is popular iff for any feasible matching M' in G, the set $\tilde{\mathcal{Y}}_{M \oplus M'}$ does not contain any of the following:*

1. *An alternating cycle with a $(1, 1)$ edge,*
2. *An alternating path which has a $(1, 1)$ edge and starts with an unmatched resident in M or a hospital which is under-subscribed in M.*
3. *An alternating path which has both its ends matched in M and has two or more $(1, 1)$ edges.*

Proof. We show that if M is a feasible matching such that for any M' the set $\tilde{\mathcal{Y}}_{M \oplus M'}$ does not contain (1), (2), (3) as in Theorem 1, then M is popular in G.

Assume for the sake of contradiction that M satisfies the conditions of Theorem 1, and yet M is not popular. Therefore there exists a feasible matching M^* such that $\Delta(M^*, M) > 0$. Consider the set $\mathcal{Y}_{M^* \oplus M}$. Recall that this set is a collection of paths and cycles and the edges of $M^* \setminus M$ are labeled. Let ρ be any path or cycle in $\mathcal{Y}_{M^* \oplus M}$ and let $\Delta(M^*, M)_\rho$ denote the difference between the votes of M^* and M when restricted to the residents and hospitals in ρ. Since $\Delta(M^*, M) > 0$, there exists a ρ such that $\Delta(M^*, M)_\rho > 0$. Note that ρ is present in $\mathcal{Y}_{M^* \oplus M}$; using the presence of ρ we establish the existence of a $\rho' \in \tilde{\mathcal{Y}}_{M^* \oplus M}$ of the form (1), (2) or (3) which contradicts our assumption. We consider three cases depending on the structure of ρ.

1. **ρ is an alternating cycle or ρ is an alternating path which starts and ends in an M edge:** Since $\rho \in \mathcal{Y}_{M^* \oplus M}$, and $\Delta(M^*, M)_\rho > 0$, it implies that there are more edges in ρ labeled $(1, 1)$ than the number of edges labeled $(-1, -1)$. We now delete the edges labeled $(-1, -1)$ from ρ; this breaks ρ in to multiple alternating paths. Note that each of these paths (say ρ') start and end with an M edge and are also present in $\tilde{\mathcal{Y}}_{M^* \oplus M}$. Furthermore, since ρ contained more number of edges labeled $(1, 1)$ than the number of edges labeled $(-1, -1)$, it is clear that there exists at least one ρ' which has two edges labeled $(1, 1)$. This is a path of type (3) from the theorem statement and therefore contradicts our assumption that M satisfied the conditions of the theorem.
2. **ρ is an alternating path which starts or ends in an M^* edge:** The proof is similar to the previous case except that when we delete from ρ the edges labeled $(-1, -1)$ we get paths $\rho' \in \tilde{\mathcal{Y}}_{M^* \oplus M}$ which are paths of type (2) or type (3) from the theorem statement. This contradicts the assumption that M satisfied the conditions of the theorem.

This completes the proof of one direction of the Theorem. To prove the other direction, we prove the contrapositive of the statement. That is, if for any feasible matching M', $\tilde{\mathcal{Y}}_{M \oplus M'}$ contains (1), (2) or (3), then M is not popular in G. We first assume that $\rho \in \tilde{\mathcal{Y}}_{M \oplus M'}$ satisfying (1), (2), or (3) is also present in $\mathcal{Y}_{M \oplus M'}$. Under this condition, it is possible to get a more popular matching than M by the following three cases.

– Let $M_2 = M \oplus \rho$ be a matching in G; by Lemma 1 we know that M_2 is feasible in G. Comparing M_2 to M yields two more votes for M_2. Hence, M_2 is more popular than M.
– If ρ is an alternating path in $\tilde{\mathcal{Y}}_{M\oplus M'}$, which has both its endpoints matched in M, and contains more than one edge labeled $(1,1)$. Then similar to the case above $M_2 = M \oplus \rho$ is more popular than M.
– If ρ is an alternating path in $\tilde{\mathcal{Y}}_{M\oplus M'}$, which has exactly one of its endpoints matched in M, and contains an edge labeled $(1,1)$, then again $M_2 = M \oplus \rho$ is more popular than M.

Now let us assume that $\rho \in \tilde{\mathcal{Y}}_{M\oplus M'}$ is not present in $\mathcal{Y}_{M\oplus M'}$. In such a case, ρ is contained in a larger path or a cycle $\rho' \in \mathcal{Y}_{M\oplus M'}$ obtained by combining ρ with other paths in $\tilde{\mathcal{Y}}_{M\oplus M'}$ and adding the deleted $(-1,-1)$ edges. Using the larger path or cycle ρ' we can construct a matching that is more popular than M. Note that we need to use paths or cycles in $\mathcal{Y}_{M\oplus M'}$ to obtain another matching, since we have to ensure that the matching obtained is indeed feasible in the instance and the correspondences are maintained. ◻

We now prove that every stable matching in an LCSM$^+$ instance is popular.

Theorem 2. *Every stable matching in an LCSM$^+$ instance G is popular.*

Proof. Let M be a stable matching in G. For any feasible matching M' in G consider the set $\mathcal{Y}_{M\oplus M'}$. To prove that M is stable, it suffices to show that there does not exist a path or cycle $\rho \in \mathcal{Y}_{M\oplus M'}$ such that an edge of ρ is labeled $(1,1)$. For the sake of contradiction, assume that ρ is such a path or cycle, which has an edge $(r', h) \in M' \setminus M$ labeled $(1,1)$. Let $r = \mathbf{corr}(r')$, where $(r,h) \in M \cap \rho$. From the proof of Lemma 1 we observe that $(M(h) \setminus \{r\}) \cup \{r'\}$ is feasible for h, therefore the edge (r', h) blocks M contradicting the stability of M. ◻

4 Popular matchings in LCSM$^+$ problem

In this section we present efficient algorithms for computing (i) a maximum cardinality popular matching, and (ii) a matching that is popular amongst all the maximum cardinality matchings in a given LCSM$^+$ instance. Our algorithms are inspired by the reductions of Kavitha and Cseh [5] where they work with a stable marriage instance. We describe a general reduction from an LCSM$^+$ instance G to another LCSM$^+$ instance G_s. Here $s = 2, \ldots, |\mathcal{R}|$. The algorithms for the two problems are obtained by choosing an appropriate value of s.
The graph G_s: Let $G = (\mathcal{R} \cup \mathcal{H}, E)$ be the input LCSM$^+$ instance. The graph $G_s = (\mathcal{R}_s \cup \mathcal{H}_s, E_s)$ is constructed as follows: Corresponding to every resident $r \in \mathcal{R}$, we have s copies of r, call them $r^0, \ldots, r^{s-1}$ in $\mathcal{R}_s$. The hospitals in $\mathcal{H}$ and their capacities remain unchanged; however we have additional dummy hospitals each of capacity 1. Corresponding to every resident $r \in \mathcal{R}$, we have $(s-1)$ dummy hospitals $d_r^0, \ldots, d_r^{s-2}$ in $\mathcal{H}_s$. Thus,

$$\mathcal{R}_s = \{\ r^0, \ldots, r^{s-1} \ \mid \forall r \in \mathcal{R}\};\quad \mathcal{H}_s = \mathcal{H} \cup \{\ d_r^0, \ldots, d_r^{s-2} \ \mid \forall r \in \mathcal{R}\}$$

We use the term level-i resident for a resident $r^i \in \mathcal{R}_s$ for $0 \le i \le s-1$. The preference lists corresponding to s different residents of r in G_s are:

– For a level-0 resident r^0, its preference list in G_s is the preference list of r in G, followed by the dummy hospital d_r^0.
– For a level-i resident r^i, where $1 \le i \le s-2$, its preference list in G_s is d_r^{i-1} followed by preference list of r in G, followed by d_r^i.
– For a level-$(s-1)$ resident r^{s-1}, its preference list in G_s is the dummy hospital d_r^{s-2} followed by the preference list of r in G.

The preference lists of hospitals in G_s are as follows.

– The preference list for a dummy hospital d_r^i is r^i followed by r^{i+1}.
– For $h \in \mathcal{H}$, its preference list in G_s, has level-$(s-1)$ residents followed by level-$(s-2)$ residents, so on upto the level-0 residents in the same order as in h's preference list in G.

Finally, we need to specify the classifications of the hospitals in G_s. For every class C_i^h in the instance G, we have a corresponding class $\bar{C}_i^h = \bigcup_{r \in C_i^h} \{r^0, \ldots, r^{s-1}\}$ in G_s, such that $q(\bar{C}_i^h) = q(C_i^h)$. We note that $|\bar{C}_i^h| = s \cdot |C_i^h|$. Let M_s be a stable matching in G_s. Then M_s satisfies the following properties:

($\mathcal{I}_1$) Each $d_r^i \in \mathcal{H}_s$ for $0 \leq i \leq s-2$, is matched to one of $\{r^i, r^{i+1}\}$ in M_s.

($\mathcal{I}_2$) The above invariant implies that for every $r \in \mathcal{R}$ at most one of $\{r^0, \ldots, r^{s-1}\}$ is assigned to a non-dummy hospital in M_s.

($\mathcal{I}_3$) For a resident $r \in \mathcal{R}$, if r^i is matched to a non-dummy hospital in M_s, then for all $0 \leq j \leq i-1$, $M_s(r^j) = d_r^j$. Furthermore, for all $i+1 \leq p \leq s-1$, $M_s(r^p) = d_r^{p-1}$. This also implies that in M_s all residents $r^0, \ldots, r^{s-2}$ are matched and only r^{s-1} can be left unmatched in M_s.

These invariants allow us to naturally map the stable matching M_s to a feasible matching M in G. We define a function $map(M_s)$ as follows.

$$M = map(M_s) = \{(r, h) : h \in \mathcal{H} \text{ and } (r^i, h) \in M_s \text{ for exactly one of } 0 \leq i \leq s-1\}$$

We outline an algorithm that computes a feasible matching in an LCSM$^+$ instance G. Given G and s, construct the graph G_s from G. Compute a stable matching M_s in G_s. If G is an LCSM$^+$ instance we use the algorithm of Huang [10] to compute a stable matching in G. If G is a PCSM$^+$ instance, it is easy to observe that G_s is also a PCSM$^+$ instance. In that case, we use the algorithm of Abraham et al.[2] to compute a stable matching. (The SPA instance is different from a PCSM$^+$ instance, however, there is a easy reduction from the PCSM$^+$ instance to SPA, we give the reduction (refer Appendix A.1) for the sake of completeness.). We output $M = map(M_s)$ whose feasibility is guaranteed by the invariants mentioned earlier. The complexity of our algorithm depends on s and the time required to compute a stable matching in the problem instance.

In the rest of the paper, we denote by M the matching obtained as $map(M_s)$ where M_s is a stable matching in G_s. For any resident $r_i \in \mathcal{R}$, we define

$$\begin{aligned} map^{-1}(r_i, M_s) &= r_i^{j_i} && \text{where } 0 \leq j_i \leq s-1 \text{ and } M_s(r_i^{j_i}) \text{ is a non-dummy hospital} \\ &= r_i^{s-1} && \text{otherwise.} \end{aligned}$$

Recall by Invariant ($\mathcal{I}_3$), exactly one of the level copy of r_i in G_s is matched to a non-dummy hospital in M_s. For any feasible matching M' in G consider the set $\mathcal{Y}_{M \oplus M'}$ – recall that this is a collection of M alternating paths and cycles in G. For any path or cycle ρ in $\mathcal{Y}_{M \oplus M'}$, let us denote by $\rho_s = map^{-1}(\rho, M_s)$ the path or cycle in G_s obtained by replacing every resident r in ρ by $map^{-1}(r, M_s)$. Recall that if a resident r is present in the class C_j^h defined by a hospital h in G, then in the graph G_s, $r^i \in \bar{C}_j^h$ for $i = 0, \ldots, s-1$. The map^{-1} function maps a resident r in G to a unique level-i copy in G_s. Using Lemma 1 and these observations we get the following corollary.

Corollary 1. *Let ρ be an alternating path or an alternating cycle in $\mathcal{Y}_{M \oplus M'}$, then $M_s \oplus \rho_s$ is a feasible matching in G_s, where $\rho_s = map^{-1}(\rho, M_s)$.*

The following technical lemma is useful in proving the properties of the matchings produced by our algorithms.

Lemma 2. *Let ρ be an alternating path or an alternating cycle in $\mathcal{Y}_{M \oplus M'}$, and $\rho_s = map^{-1}(\rho, M_s)$.*

1. *There cannot be any edge labeled $(1,1)$ in ρ_s.*
2. *Let $\langle r_a^{j_a}, h, r_b^{j_b} \rangle$ be a sub-path of ρ_s, where $h = M_s(r_b^{j_b})$. Then, the edge $(r_a^{j_a'}, h) \notin \rho_s$ cannot be labeled $(1,1)$, where $j_a' < j_a$.*

Proof. Let $\langle r_a^{j_a}, h, r_b^{j_b} \rangle$ be a sub-path of ρ_s, where $h = M_s(r_b^{j_b})$ (Figure 4). As $M_s \oplus \rho_s$ is feasible in G_s (Corollary 1), the set $(M_s(h) \setminus \{r_b^{j_b}\}) \cup \{r_a^{j_a}\}$ is feasible for h in G_s. Now since $(r_a^{j_a}, h)$ is labeled $(1,1)$, the edge $(r_a^{j_a}, h)$ blocks M_s contradicting its stability. This proves (1). To prove (2), assume that the edge $(r_a^{j_a'}, h) \notin \rho_s$ is labeled $(1,1)$. The residents $r_a^{j_a}$ and $r_a^{j_a'}$ belong to the same class (say $\bar{C}_k^h$) in G_s, hence $(M_s(h) \setminus \{r_b^{j_b}\}) \cup \{r_a^{j_a'}\}$ is feasible for h. Thus the edge $(r_a^{j_a'}, h)$ blocks M_s contradicting its stability.

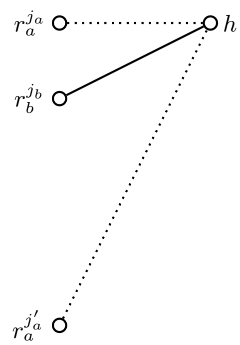


Fig. 4. The edges $(r_a^{j_a}, h)$ and $(r_b^{j_b}, h)$ belong to ρ, while the edge $(r_a^{j'_a}, h)$ does not belong to ρ.

4.1 Maximum cardinality popular matching

Let $G = (\mathcal{R} \cup \mathcal{H}, E)$ be an instance of the LCSM$^+$ problem where we are interested in computing a maximum cardinality popular matching. We use our generic reduction with the value of the parameter $s = 2$. Since G_2 is linear in the size of G, and a stable matching in an LCSM$^+$ instance can be computed in $O(mn)$ time [10], we obtain an $O(mn)$ time algorithm to compute a maximum cardinality popular matching in G. In case G is a PCSM$^+$ instance, we use the linear time algorithm in [2] for computing a stable matching to get a linear time algorithm for our problem. The proof of correctness involves two things – we first show that M is popular in G. We then argue that it is the largest size popular matching in G. We state the main theorem of this section below.

Theorem 3. *Let $M = map(M_2)$ where M_2 is a stable matching in G_2. Then M is a maximum cardinality popular matching in G.*

We break down the proof of Theorem 3 in two parts. Lemma 3 shows that the assignment M satisfies all the conditions of Theorem 1. Lemma 5 shows that the matching output is indeed the largest size popular matching in the instance. Let M' be any assignment in G. Recall the definition of $\tilde{\mathcal{Y}}_{M \oplus M'}$ – this set contains M alternating paths and M alternating cycles in G and the edge labels on the M' edges belong to $\{(-1,1), (1,-1), (1,1)\}$.

Lemma 3. *Let $M = map(M_2)$ where M_2 is a stable matching in G_2 and let M' be any feasible assignment in G. Consider the set of alternating paths and alternating cycles $\tilde{\mathcal{Y}}_{M \oplus M'}$. Then, the following hold:*

1. *An alternating cycle C in $\tilde{\mathcal{Y}}_{M \oplus M'}$, does not contain any edge labeled $(1,1)$.*
2. *An alternating path P in $\tilde{\mathcal{Y}}_{M \oplus M'}$ that starts or ends with an edge in M', does not contain any edge labeled $(1,1)$.*
3. *An alternating path P in $\tilde{\mathcal{Y}}_{M \oplus M'}$ which starts and ends with an edge in M, contains at most one edge labeled $(1,1)$.*

Proof. We first prove the parts (1) and (2). Recall that $M = map(M_2)$ where M_2 is a stable matching in G_2. Assume that $\rho = \langle u_0, v_1, u_1, \ldots, v_k, u_k \rangle$ where for each $i = 0, \ldots, k$, $v_i = M(u_i)$ (in case u_i is a hospital, $v_i \in M(u_i)$). In case ρ is a cycle, all subscripts follow mod k arithmetic. The existence of ρ in $\tilde{\mathcal{Y}}_{M \oplus M'}$ implies that there is an associated M_2 alternating path or an M_2 alternating cycle $\rho_2 = map^{-1}(\rho, M_2)$ in G_2.

Now assume for the sake of contradiction that ρ contains an edge $e = (r_a, h_b) \notin M$ labeled $(1,1)$ for some $a = 0, \ldots, k$, and $b = 0, \ldots, k$. We observe the following about preferences of r_a and h_b in G.

($\mathcal{O}_1$) r_a prefers h_b over $h_a = M(r_a)$.
($\mathcal{O}_2$) h_b prefers r_a over $r_b \in M(h_b)$, where $r_b = \mathbf{corr}(r_a)$.

Using the presence of an edge labeled $(1,1)$ in ρ, we will contradict the stability of M_2 in G_2. Consider the edge $e' = (r_a^{j_a}, h_b)$ in G_2. Since $h_b = M_2(r_b^{j_b})$, we observe that $e' \notin M_2$. We consider the four cases that can arise depending on the values of j_a and j_b.

1. $j_a = j_b = 0$
2. $j_a = j_b = 1$
3. $j_a = 1$ and $j_b = 0$
4. $j_a = 0$ and $j_b = 1$

Recall observation $(\mathcal{O}_1)$, and the fact that the residents do not change their preferences in G_2 w.r.t. the hospitals originally in G. This implies in all the four cases above, the resident $r_a^{j_a}$ prefers h_b over $h_a = M_2(r_a^{j_a})$. Using $(\mathcal{O}_2)$ and the fact that a hospital h in G_2 prefers level-1 residents over level-0 residents, we can conclude the following. For the cases $(1), (2)$ and (3), hospital h_b prefers $r_a^{j_a}$ over $r_b^{j_b}$, which implies that the pair $(r_a^{j_a}, h_b)$ is labeled $(1,1)$, and thus forms a blocking pair w.r.t. M_2 (using Lemma 2(1)).

We now consider the three different cases for ρ depending on whether ρ is a path or a cycle. When ρ is a path, we break down its proof in two cases – (i) ρ starts or ends with a resident unmatched in M. (ii) ρ starts or ends with an under-subscribed hospital. In each of the different possibilities for ρ, we show that the stability of M_2 can be contradicted even in case (4), i.e. when $j_a = 0$ and $j_b = 1$.

- $\rho = \langle r_0, h_1, r_1, \ldots, h_{k-1}, r_{k-1}\rangle$ is an alternating path that starts or ends with a resident which is unmatched in M. Here $\rho_2 = map^{-1}(\rho, M_2) = \langle r_0^{j_0}, h_1, r_1^{j_1}, \ldots, h_{k-1}, r_{k-1}^{j_{k-1}}\rangle$ and for $t = 0, \ldots, k-1$, $j_t \in \{0,1\}$.
  Using invariants $(\mathcal{I}_1)$, $(\mathcal{I}_2)$, and $(\mathcal{I}_3)$, we conclude that a resident r remains unmatched in M_2 when its level-0 copy is matched to the dummy hospital d_r, and the level-1 copy is unmatched in M_2. Therefore, the first resident on the path ρ_2 is a level-1 resident. Furthermore, the second resident on the path r_1 has to be a level-1 resident. Otherwise, as r_0^1 is unmatched in M_2 and h_1 prefers a level-1 resident over a level-0 resident, the edge (r_0^1, h_1) will be labeled $(1,1)$, and thus forms a blocking pair w.r.t. M_2 (using Lemma 2(1)).
  We consider an edge $e \in \rho$ such that $b = a+1$. In case (4), we observe that as $j_0 = j_1 = 1, j_a = 0$, and $a < b$, there exists an index x in ρ_2 such that there is a transition from a level-1 resident to a level-0 resident. That is, $(r_x^0, h_x) \in M_2$ and $(r_{x-1}^1, h_x) \notin M_2$ both belong to ρ_2.
  We enumerate the possible labels for the edge $e_x = (r_{x-1}, h_x)$ in G.
  - If e_x is labeled $(1,1)$ or $(1,-1)$, then the edge (r_{x-1}^1, h_x) is labeled $(1,1)$, and thus blocks M_2 (using Lemma 2(1)).
  - If e_x is labeled $(-1,1)$, then the edge (r_{x-1}^0, h_x) is labeled $(1,1)$, and thus blocks M_2 (using Lemma 2(2)).
- $\rho = \langle h_0, r_1, h_1, \ldots, r_{k-1}, h_{k-1}\rangle$ is an alternating path that starts or ends with an under-subscribed hospital. Here $\rho_2 = map^{-1}(\rho, M_2) = \langle h_0, r_1^{j_1}, h_1, \ldots, r_{k-1}^{j_{k-1}}, h_{k-1}\rangle$ and for $t = 1, \ldots, k-1$, $j_t \in \{0,1\}$.
  Observe that if $j_1 = 1$, then (r_1^0, h_0) is labeled $(1,1)$, as h_0 is unmatched in M, and r_1^0 prefers h_0 to d_{r_1} ($d_{r_1} = M_2(r_1^0)$ using invariants $(\mathcal{I}_1)$, $(\mathcal{I}_2)$, and $(\mathcal{I}_3)$), contradicting the stability of M_s (using Lemma 2(1)). Thus, it must be the case that $j_0 = 0$. Note that the edge (r_1, h_0) can not be labeled $(1,1)$ in G, as h_0 being under-subscribed prefers being matched to r_1, and residents do not change their votes, and thus the edge (r_1^0, h_0) is labeled $(1,1)$, contradicting the stability of M_2 (using Lemma 2(1)).
  We consider an edge $e \in \rho$ such that $a = b+1$. In case (4), we observe that as $j_1 = 0, j_b = 1$, and $a > b$, there exists an index x in ρ_2 such that there is a transition from a level-0 resident to a level-1 resident. That is, $(r_x^0, h_x) \in M_2$ and $(r_{x+1}^1, h_x) \notin M_2$ both belong to ρ_2. Using an argument similar to in the case above, we can show that either the edge (r_{x+1}^1, h_x) or the edge (r_{x+1}^0, h_x) is labeled $(1,1)$, and therefore forms a blocking pair w.r.t. M_2.
- $\rho = \langle r_0, h_0, r_1, h_1, \ldots, r_k, h_k, r_0\rangle$ is an alternating cycle. Here $\rho_2 = map^{-1}(\rho, M_s) = \langle r_0^{j_0}, h_0, r_1^{j_1}, h_1, \ldots, r_k^{j_k}, h_k, r_0^{j_0}\rangle$ and for $t = 1, \ldots, k-1$, $j_t \in \{0,1\}$.
  We consider an edge $e \in \rho$ such that $a = b+1$. As $j_a = 0$ and $j_b = 1$, and $b < a$, this is a transition from a level-1 resident to a level-0 resident in the cycle ρ_2. To complete the cycle ρ_2 there must exist an index x such that there is a transition from a level-0 resident to a level-1 resident. That is, $(r_x^0, h_x) \in M_2$ and $(r_{x+1}^1, h_x) \notin M_2$ both belong to ρ_2. Using an argument similar to as in the first case, we can show that either the edge (r_{x+1}^1, h_x) or the edge (r_{x+1}^0, h_x) is labeled $(1,1)$, and therefore forms a blocking pair w.r.t. M_2.

We now prove part (3) of the lemma. Consider $P = \langle r_0, h_0, \ldots, r_{k-1}, h_{k-1}\rangle$ where for each $i = 0, \ldots, k-1$, $M(r_i) = h_i$. The existence of P in $\tilde{\mathcal{Y}}_{M\oplus M'}$ implies that there exists an M_2 alternating path P_2 in G_2. Here $P_2 = \langle r_0^{j_0}, h_0, \ldots, r_{k-1}^{j_{k-1}}, h_{k-1}\rangle$, and for $t = 0, \ldots, k-1$, $j_t \in \{0, 1\}$.

For the sake of contradiction assume that P contains at least two edges, $e_1 = (r_x, h_{x-1})$, $e_2 = (r_y, h_{y-1})$ for some $x, y = 1, \ldots, k-1$, w.l.o.g. $x \neq y, x < y$ and e_1, e_2 are labeled $(1, 1)$. We observe the following about preferences of r_x, r_y and h_{x-1}, h_{y-1} in G.

($\mathcal{O}_1$) r_x prefers h_{x-1} over $h_x = M(r_x)$.
r_y prefers h_{y-1} over $h_y = M(r_y)$.
($\mathcal{O}_2$) h_{x-1} prefers r_x over $r_{x-1} \in M(h_{x-1})$.
h_{y-1} prefers r_y over $r_{y-1} \in M(h_{y-1})$.

Using the presence of the edges e_1 and e_2 labeled $(1, 1)$ in P, we will contradict the stability of M_2 in G_2. Consider the edges $e_1' = (r_x^{j_x}, h_{x-1})$ and $e_2' = (r_y^{j_y}, h_{y-1})$ in G_2, and since $h_{x-1} = M_2(r_x^{j_{x-1}})$ and $h_{y-1} = M_2(r_y^{j_{y-1}})$, note that $e_1', e_2' \notin M_2$.

We first consider the edge e_1', and consider the four cases that can arise depending on the values of j_x and j_{x-1}.

1. $j_{x-1} = j_x = 0$
2. $j_{x-1} = j_x = 1$
3. $j_{x-1} = 0$ and $j_x = 1$
4. $j_{x-1} = 1$ and $j_x = 0$

Recall observation ($\mathcal{O}_1$), and the fact that the residents do not change their preferences in G_2 w.r.t the hospitals originally in G. This implies that in all the four cases above, the resident $r_x^{j_x}$ prefers h_{x-1} over $h_x = M_2(r_x^{j_x})$. Using ($\mathcal{O}_2$) and the fact that a hospital h in G_2 prefers level-1 residents over level-0 residents, we can conclude the following. For the cases (1), (2) and (3), hospital h_{x-1} prefers $r_x^{j_x}$ over $r_{x-1}^{j_{x-1}}$, which implies that the pair $(r_x^{j_x}, h_{x-1})$ is labeled $(1, 1)$, which contradicts the stability of M_s (using Lemma 2(1)).

With a similar analysis for the edge e_2', we conclude that the first three cases do not arise. There is only one case left to consider, when $j_{x-1} = 1, j_x = 0$ and $j_{y-1} = 1, j_y = 0$. As $x \neq y, x < y$, and $j_x = 0$, $j_{y-1} = 1$, there exists an index ℓ in P_2 such that there is a transition from a level-0 resident to a level-1 resident. That is, $(r_\ell^0, h_\ell) \in M_2$ and $(r_{\ell+1}^1, h_\ell) \notin M_2$ both belong to P_2.

We enumerate the possible labels for the edge $e_\ell = (r_{\ell+1}, h_\ell)$ in G.

- If e_ℓ is labeled $(1, 1)$ or $(1, -1)$, then the edge $(r_{\ell+1}^1, h_\ell)$ is labeled $(1, 1)$, which contradicts the stability of M_s (using Lemma 2(1)).
- If e_ℓ is labeled $(-1, 1)$, then the edge $(r_{\ell+1}^0, h_\ell)$ is labeled $(1, 1)$, which contradicts the stability of M_s (using Lemma 2(2)).

This completes the proof. □

Lemma 4. *There is no augmenting path with respect to M in $\tilde{\mathcal{Y}}_{M\oplus M'}$.*

Proof. Let $P = \langle r_0, h_1, r_1, h_2, \ldots, h_{k-1}, r_{k-1}, h_k\rangle$ be an augmenting path where for each $i = 1, \ldots, k-1$, $M(r_i) = h_i$. The existence of P in $\tilde{\mathcal{Y}}_{M\oplus M'}$ implies that there exists an M_2 augmenting path $P_2 = \langle r_0^{j_0}, h_1, r_1^{j_1}, h_2, \ldots, h_{k-1}, r_{k-1}^{j_{k-1}}, h_k\rangle$ in G_2, and for $t = 0, \ldots, k-1$, $j_t \in \{0, 1\}$.

Using invariants ($\mathcal{I}_1$), ($\mathcal{I}_2$), and ($\mathcal{I}_3$), we conclude that a resident r remains unmatched in M_2 when its level-0 copy is matched to the dummy vertex d_r, and the level-1 copy is unmatched in M_2. Therefore the first resident on the path P_2 is a level-1 resident. The second resident on the path r_1 has to be a level-1 resident, otherwise the edge (r_0^1, h_1) will be labeled $(1, 1)$, and thus contradict the stability of M_2 (using Lemma 2(1)). This is because r_0^1 prefers being matched to h_1 than being unmatched in M_2, and h_1 prefers level-1 resident over a level-0 resident. Observe that $j_{k-1} = 0$, else the pair (r_{k-1}^0, h_k) is labeled $(1, 1)$, as r_{k-1}^0 is matched to d_r (by invariants ($\mathcal{I}_1$) and ($\mathcal{I}_2$)), which is at the end of its preference list, and h_k is unmatched in M'.

Therefore the path P_2 is of the form $\langle r_0^1, h_1, r_1^1, h_2, \ldots, h_{k-1}, r_{k-1}^0, h_k\rangle$. As $j_0 = j_1 = 1$ and $j_{k-1} = 0$, there exists an index x in P_2 such that there is a transition from a level-1 resident to a level-0 resident. That is, $(r_x^0, h_x) \in M_2$ and $(r_{x-1}^1, h_x) \notin M_2$ both belong to P_2.

We enumerate the possible labels for the edge $e_x = (r_{x-1}, h_x)$ in G.

– If e_x is labeled $(1, 1)$ or $(1, -1)$, then the edge (r^1_{x-1}, h_x) is labeled $(1, 1)$, which contradicts the stability of M_s (using Lemma 2(1)).
– If e_x is labeled $(-1, 1)$, then the edge (r^0_{x-1}, h_x) is labeled $(1, 1)$, which contradicts the stability of M_s (using Lemma 2(2)).

This contradicts our assumption that P is augmenting with respect to M in $\tilde{\mathcal{Y}}_{M\oplus M'}$. □

Lemma 5. *There exists no popular matching M^* in G such that $|M^*| > |M|$.*

Proof. For contradiction, assume that such an assignment M^* exists in G. Consider the set $\mathcal{Y}_{M\oplus M^*}$; recall that this set contains alternating paths and cycles possibly containing edges labeled $(-1, -1)$. Since $|M^*| > |M|$ there must exist an augmenting path P in $\mathcal{Y}_{M\oplus M^*}$. We first claim that the path P must contain at least one edge labeled $(-1, -1)$. If not, then the path P is also contained in $\tilde{\mathcal{Y}}_{M\oplus M^*}$. However, by Lemma 4 there is no augmenting path with respect to M in $\tilde{\mathcal{Y}}_{M\oplus M'}$ for any feasible matching M' in G.

We now remove all edges from P which are labeled $(-1, -1)$. This breaks the path into sub-paths say $P_1, P_2, \ldots, P_t$ for some $t \geq 1$, where P_1 and P_t have one endpoint unmatched in M. Consider the path P_1; since P_1 does not contain any $(-1, -1)$ edge this implies that $P_1 \in \tilde{\mathcal{Y}}_{M\oplus M^*}$. Without loss of generality, assume that P_1 starts with a resident r which is unmatched in M. Thus using Lemma 3(2), P_1 does not contain any edge labeled $(1, 1)$. Let us denote by $\Delta(M^*, M)_{P_1}$ the difference between votes of M^* and M restricted to vertices of path P_1. It is clear that $\Delta(M^*, M)_{P_1} < 0$. Also, for each $i = 2, \ldots, t-1$, the alternating paths P_i have both of their endpoints matched in M. Thus we have $\Delta(M^*, M)_{P_i} \leq 0$ as there can be at most one $(1, 1)$ edge (by Lemma 3(3)) in these paths, but the endpoints prefer M, as they are matched in M but not in M'. If P_t exists, then a argument similar as given for P_1, we have $\Delta(M^*, M)_{P_t} < 0$. Using these observations, we conclude that M is more popular than M^*, a contradiction to the assumption that M^* and M are both popular.

Thus, for any given matching M^* such that $|M^*| > |M|$, we know that M is more popular than such a matching. This completes the proof of the lemma, and shows that the matching $M = map(M_2)$ is a maximum cardinality popular matching in G. □

4.2 Popular matching amongst maximum cardinality matchings

In this section we give an efficient algorithm for computing a matching which is *popular* amongst the set of maximum cardinality matchings. The matching M that we output cannot be beaten in terms of votes by any feasible maximum cardinality matching. Our algorithm uses the generic reduction with a value of $s = |\mathcal{R}| = n_1$ (say). Thus, $|\mathcal{R}_{n_1}| = n_1^2$, and $|\mathcal{H}_{n_1}| = |\mathcal{H}| + O(n_1^2)$. Furthermore, $|E_{n_1}| = O(mn_1)$ where $m = |E|$. Thus the running time of the generic algorithm presented earlier with $s = n_1$ for an LCSM$^+$ instance is $O(mn \cdot n_1) = O(mn^2)$ and for a PCSM$^+$ instance is $O(mn_1) = O(mn)$.

To prove correctness, we show that the matching output by our algorithm is (i) maximum cardinality and (ii) popular amongst all maximum cardinality feasible matchings. Let $M = map(M_{n_1})$ and M^* be any maximum cardinality feasible matching in G. Consider the set $\mathcal{Y}_{M\oplus M^*}$, and let ρ be an alternating path or an alternating cycle in $\mathcal{Y}_{M\oplus M^*}$. Let $\rho_{n_1} = map^{-1}(\rho, M_{n_1})$ denote the associated alternating path or cycle in G_{n_1}. We observe that every hospital on the path ρ_{n_1} is a non-dummy hospital since ρ_{n_1} was obtained using the inverse-map of ρ. We observe two useful properties about such a path or cycle ρ_{n_1} in G_{n_1}. We show that if for a hospital $h \in \rho_{n_1}$, the level of the unmatched resident incident on h is greater than the level of the matched resident incident on h, then such a level change is *gradual*, and the associated edge in ρ has the label $(-1, -1)$. Lemma 6, gives a proof of these.

Lemma 6. *Let ρ_{n_1} be an alternating path or an alternating cycle in G_{n_1} and let h be a hospital which has degree two in ρ_{n_1}. Let $\langle r_a^{j_a}, h, r_b^{j_b}\rangle$ be the sub-path containing h where $M(r_b^{j_b}) = h$. If $j_a > j_b$, we claim the following:*

1. *$j_a = j_b + 1$.*
2. *The associated edge $(r_a, h) \in \rho$ is labeled $(-1, -1)$.*

Proof. We first prove that $j_a = j_b + 1$. For contradiction, assume that $j_a > j_b + 1$. Observe that h prefers all the level-j_a residents over any level-j_b resident. We consider the edge $e = (r_a, h)$ in the graph G. We claim that the label for the edge e cannot be $(1,1)$ or $(1,-1)$, otherwise the edge $(r_a^{j_a}, h)$ is labeled $(1,1)$ in G_{n_1} as the residents do not change their votes. Similarly, we claim that the label for the edge e cannot be $(-1,1)$ or $(-1,-1)$, as $r_a^{j_a-1}$ is matched in M_{n_1} to the last dummy on its preference list, $d_{r_a}^{j_a-1} = M_{n_1}(r_a^{j_a-1})$ (by invariant $(\mathcal{I}_3)$), and prefers h to $d_{r_a}^{j_a-1}$, and h prefers all the level-(j_a-1) residents over any level-j_b resident. In this case the edge $(r_a^{j_a-1}, h)$ is labeled $(1,1)$ in G_{n_1}, and thus blocks M_{n_1} (by Lemma 2(2)).

To prove part (b), we assume $j_a = j_b + 1$. We enumerate the possible labels for the edge $e = (r_a, h)$ in G.

– If e is labeled $(1,1)$ or $(1,-1)$, then the edge $(r_a^{j_a}, h)$ is labeled $(1,1)$, as $r_a^{j_a}$ prefers h over $M_{n_1}(r_a^{j_a})$, and h prefers any level-j_a resident over a level-j_b resident. Thus, the edge $(r_a^{j_a}, h)$ blocks M_{n_1} (by Lemma 2(1)).
– If e is labeled $(-1,1)$, then the edge $(r_a^{j_b}, h)$ is labeled $(1,1)$, as $r_a^{j_b}$ prefers h over $d_{r_a}^{j_b} = M_{n_1}(r_a^{j_b})$, and h prefers $r_a^{j_b}$ over $r_b^{j_b}$ according to its preference list. Thus, the edge $(r_a^{j_b}, h)$ blocks M_{n_1} (by Lemma 2(2)).

Thus, the only possible label for the edge (r_a, h) is $(-1,-1)$. □

We use Lemma 7 to prove that M is a maximum cardinality matching in G.

Lemma 7. *Let M^* be any feasible maximum cardinality matching in G. Then there is no augmenting path with respect to M in $\mathcal{Y}_{M \oplus M^*}$.*

Proof. For the sake of contradiction assume that the path $P = \langle r_0, h_1, r_1, \ldots, h_{k-1}, r_{k-1}, h_k \rangle$ is an augmenting path where for each $i = 1, \ldots, (k-1)$, $M(r_i) = h_i$. Here r_0 is unmatched in M, and h_k is under-subscribed in M. The existence of P in $\mathcal{Y}_{M \oplus M^*}$ implies that there exists an M_{n_1} augmenting path $P_{n_1} = map^{-1}(P, M_{n_1}) = \langle r_0^{j_0}, h_1, r_1^{j_1}, \ldots, h_{k-1}, r_{k-1}^{j_{k-1}}, h_k \rangle$ in G_{n_1}, and for $t = 0, \ldots, k-1$, $j_t \in \{0, \ldots, n_1 - 1\}$, where $h_i = M_{n_1}(r_i^{j_i})$.

Since r_0 is unmatched in M, by invariant $(\mathcal{I}_3)$, it implies that for $0 \le i \le n_1 - 2$, $M_{n_1}(r_0^i) = d_{r_0}^i$, and $r_0^{n_1-1}$ is unmatched in M_{n_1}. This implies that the first resident in the path P_{n_1} is a level-(n_1-1) resident $r_0^{n_1-1}$. The second resident on the path P_{n_1} also has to be a level-(n_1-1) resident. If not, then the edge $(r_0^{n_1-1}, h_1)$ is labeled $(1,1)$ since h_1 prefers $r_0^{n_1-1}$ to any resident at a level lower than $n_1 - 1$ and $r_0^{n_1-1}$ is unmatched in M_{n_1}. The last resident in the path P_{n_1} is a level-0 resident i.e. $r_{k-1}^{j_{k-1}} = r_{k-1}^0$. If not, then the edge (r_{k-1}^0, h_k) is labeled $(1,1)$, as r_{k-1}^0 is matched to the last dummy hospital $(d_{r_{k-1}}^0)$ on its preference list (by invariant $(\mathcal{I}_3)$), and h_k is under-subscribed in M_{n_1}.

Thus, in the path P_{n_1}, the first two residents are level-(n_1-1), while the last resident is level-0. Recall that the path P_{n_1} was obtained as an inverse-map of the path P in G. Since the path P contains at most n_1 residents (possibly all of the residents in G), the path P_{n_1} also contains at most n_1 residents. From Lemma 6 we observe that the difference in the levels of two residents in a sub-path of P_{n_1} can be at most one. Thus, it must be the case that residents at all the levels $n_1 - 1$ to 0 are present in P_{n_1}. However, since there are two residents at level-(n_1-1) (first two residents) and one resident at level-0 (last resident), it is clear that residents at all levels from $n_1 - 1$ to 0 cannot be accommodated in a path containing at most n_1 residents.

This contradicts the existence of such a path P_{n_1} in G_{n_1} which implies that the assumed augmenting path P with respect to M cannot exist. This proves that $M = map(M_{n_1})$ is a max-cardinality matching in G. □

We can now conclude that the set $\mathcal{Y}_{M \oplus M^*}$ is a set of alternating (and not augmenting) paths and alternating cycles. It remains to show that M is popular amongst all maximum cardinality feasible matchings in G. Let M^* be any feasible maximum cardinality matching in G. In Lemma 8 we show that if there is an edge $(r, h) \in M^* \setminus M$ labeled $(1,1)$ in ρ, then in ρ_{n_1}, for the hospital h, the level of its unmatched neighbour (resident) is lower than the level of its matched neighbour (resident).

Lemma 8. *If an edge $(r_a, h) \in \rho$ is labeled $(1,1)$, then in ρ_{n_1} for the sub-path $\langle r_a^{j_a}, h, r_b^{j_b} \rangle$ where $M_{n_1}(r_b^{j_b}) = h$, we have $j_a < j_b$.*

Proof. Let an edge $e = (r_a, h)$ be labeled $(1, 1)$ in ρ. We observe the following about preferences of r_a and h in G.

($\mathcal{O}_1$) r_a prefers h over $M(r_a)$.
($\mathcal{O}_2$) h prefers r_a over $r_b \in M(h)$, where $r_b = \mathbf{corr}(r_a)$.

Consider the edge $e' = (r_a^{j_a}, h)$ in G_{n_1}, as $h = M_{n_1}(r_b^{j_b})$ it implies $e' \notin M_{n_1}$. Consider the three cases that can arise depending on the values of j_a and j_b.

1. $j_a = j_b$
2. $j_a > j_b$
3. $j_a < j_b$

Recall observation ($\mathcal{O}_1$), and the fact that the residents do not change their preferences in G_{n_1} w.r.t. the hospitals originally in G. This implies in all the three cases above, the resident $r_a^{j_a}$ prefers h over $M_{n_1}(r_a^{j_a})$. Using ($\mathcal{O}_2$) and the fact that a hospital h in G_{n_1} prefers level-p residents over level-q residents, when $p > q$, we can conclude the following. For the cases (1) and (2), hospital h prefers $r_a^{j_a}$ over $r_b^{j_b}$, which implies that the pair $(r_a^{j_a}, h)$ is labeled $(1, 1)$, which is a blocking pair for M_{n_1} (using Lemma 2(1)). This contradicts the stability of M_{n_1}. We therefore conclude that $j_a < j_b$. □

Lemma 9 shows that in an alternating path in $\mathcal{Y}_{M \oplus M^*}$ with exactly one endpoint unmatched in M or an alternating cycle, the number of edges labeled $(1, 1)$ cannot exceed the number of edges labeled $(-1, -1)$.

Lemma 9. *Let ρ be an alternating path or an alternating cycle in $\mathcal{Y}_{M \oplus M^*}$. Then the number of edges labeled $(1, 1)$ in ρ is at most the number of edges labeled $(-1, -1)$.*

Proof. Depending on the nature of ρ we have three different cases.

- ρ is an alternating path which starts with an unmatched resident in M.
- ρ is an alternating path which starts with a hospital which is under-subscribed in M.
- ρ is an alternating cycle.

The proof idea is similar in all the three cases. In each of the above, we consider $\rho_{n_1} = map^{-1}(\rho, M_{n_1})$. For every edge labeled $(1, 1)$ in ρ we show a change (increase / decrease) in the level of the residents which are neighbours of a particular hospital. We show that each such change must be complemented with another change (decrease / increase resp.) in the level of the residents which are neighbours to some other hospital. Finally, we show that the second type of change translates to a $(-1, -1)$ edge in ρ.

Let ρ be an alternating cycle. Consider a hospital $h_i \in \rho$ for which there is an edge (r, h) labeled $(1, 1)$ incident on it in ρ. Consider the associated hospital $h_i \in \rho_{n_1}$. W.l.o.g. let $\langle r_{i-1}^{j_{i-1}}, h_i, r_i^{j_i} \rangle$ be a sub-path of ρ when traversing ρ_{n_1} in counter-clock-wise direction. By Lemma 8, we know that the level of the unmatched resident incident on h_i is lower than the level of the matched resident incident on h_i. Thus there is an *increase* in level of residents when at h_i (while traversing ρ_{n_1} in counter-clockwise direction). This is true for any $h_k \in \rho_{n_1}$ where the associated hospital in ρ has a $(1, 1)$ edge incident on it. We now recall from Lemma 6(2) that whenever a hospital $h_k \in \rho_{n_1}$ has a level *decrease*, the associated edge in ρ is labeled $(-1, -1)$. Furthermore the *decrease* in levels at a hospital is gradual. Thus, it must be the case that the number of $(1, 1)$ edges in ρ is at most the number of $(-1, -1)$ edges in ρ.

In case ρ is an alternating path starting at an unmatched resident, we show that in the path ρ_{n_1} the first two residents are level-$(n_1 - 1)$ residents (see **Claim 1** below for a proof). Furthermore, consider the first edge $(r, h) \in \rho$ that is labeled $(1, 1)$. The associated hospital h_i has a *increase* in the level of its two neighbouring residents. However, since ρ_{n_1} started with two level-$(n_1 - 1)$ residents (which is the highest level possible). Therefore, there must have been some hospital h_k preceding h_i in ρ_{n_1} which has a *decrease* in the levels of the two neighbours. Using these facts it is easy to prove the following:
Number of $(1, 1)$ edges in $\rho \leq$ Number of increases in $\rho_{n_1} \leq$ Number of decreases in $\rho_{n_1} \leq$ Number of $(-1, -1)$ edges in ρ.
This completes the proof in case ρ is a path starting at an unmatched resident.

Finally, we are left with the case when ρ is an alternating path which starts with a hospital h_i which is under-subscribed in M. We show that in the associated path ρ_{n_1}, the first resident is a level-0 resident (see

Claim 2 below for a proof). Note that in this case since the path starts at a hospital, whenever we have an edge labeled $(1,1)$ in ρ, the associated hospital h_i in ρ_{n_1} has an *increase* in the levels of the two neighbouring residents. Now first resident in the path is at the lowest possible level, it must be the case that there is a hospital h_k preceding h_i in ρ for which there is a *decrease* in the level of the neighbouring residents. Now using arguments similar to those in the case of path starting at an unmatched resident, we conclude that the number of $(1,1)$ edges in ρ is at most the number of $(-1,1)$ edges in ρ.

Claim 1: $\rho = \langle r_0, h_1, r_1, \ldots \rangle$ starts with an unmatched resident. As r_0 is unmatched in M, by invariant $(\mathcal{I}_3)$, it implies that for $0 \leq i \leq n_1 - 2$, $M_{n_1}(r_0^i) = d_r^i$, and $r_0^{n_1-1}$ is unmatched in M_{n_1}. Therefore the first resident $r_0^{j_0}$ on the path ρ_{n_1} is a level-$(n_1 - 1)$ resident, that is $j_0 = n_1 - 1$. Furthermore, the second resident on the path r_1 has to be a level-$(n_1 - 1)$ resident. If not, then as $r_0^{n_1-1}$ is unmatched in M_{n_1} and h_1 prefers a level-$(n_1 - 1)$ resident to a level-v resident ($v < n_1 - 1$), the edge $(r_0^{n_1-1}, h_1)$ will be labeled $(1,1)$, and thus blocks M_{n_1} contradicting its stability (using Lemma 2(1)).

Claim 2: $\rho = \langle h_0, r_1, h_1, \ldots \rangle$ starts with an under-subscribed hospital. The first resident $r_1^{j_1}$ on the path ρ_{n_1} has to be a level-0 resident, that is $j_1 = 0$. If not, i.e. if $j_1 = 1$, then $d_{r_0}^0 = M_{n_1}(r_1^0)$ (by invariant $(\mathcal{I}_3)$) $d_{r_0}^0$ in M_{n_1}, and prefers h_1 to $d_{r_0}^0$. The hospital h_1 on the other hand is under-subscribed in M_{n_1} and prefers being matched to r_1^0 in M_{n_1}. Thus, the edge (r_1^0, h_1) is labeled $(1,1)$, and blocks M_{n_1} contradicting its stability (using Lemma 2(2)). □

Thus, we get the following theorem:

Theorem 4. *Let $M = map(M_{n_1})$ where M_{n_1} is a stable matching in G_{n_1}. Then M is a popular matching amongst all maximum cardinality matchings in G.*

Discussion: A natural question is to consider popular matchings in LCSM instances. An LCSM instance need not admit a stable matching. However we claim that restricted to LCSM instances which admit a stable matching, our results hold without any modification. To obtain the result, we claim that Lemma 1 holds in the presence of lower quotas on classes. Additionally, if the given LCSM instance G admits a stable matching, the graph G_s for $s = 1, \ldots, n_1$ also admits a stable matching. We thank Prajakta Nimbhorkar for pointing this to us.

Acknowledgement: We thank the anonymous reviewers whose comments have improved the presentation.

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. D. J. Abraham, R. W. Irving, and D. F. Manlove. Two Algorithms for the Student-Project Allocation Problem. J. of Discrete Algorithms, 5(1):73–90, 2007.
3. P. Biró, D. Manlove, and S. Mittal. Size Versus Stability in the Marriage Problem. Theoretical Computer Science, 411(16-18):1828–1841, 2010.
4. F. Brandl and T. Kavitha. Popular Matchings with Multiple Partners. CoRR, abs/1609.07531, 2016.
5. Á. Cseh and T. Kavitha. Popular Edges and Dominant Matchings. In Proceedings of the Eighteenth Conference on Integer Programming and Combinatorial Optimization, pages 138–151, 2016.
6. T. Fleiner and N. Kamiyama. A Matroid Approach to Stable Matchings with Lower Quotas. In Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms, pages 135–142, 2012.
7. D. Gale and L. Shapley. College Admissions and the Stability of Marriage. American Mathematical Monthly, 69:9–14, 1962.
8. P. Gärdenfors. Match Making: assignments based on bilateral preferences. Behavioural Sciences, 20:166–173, 1975.
9. D. Gusfield and R. W. Irving. The Stable Marriage Problem: Structure and Algorithms. MIT Press, 1989.
10. C.-C. Huang. Classified Stable Matching. In Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, pages 1235–1253, 2010.
11. C.-C. Huang and T. Kavitha. Popular Matchings in the Stable Marriage Problem. In Proceedings of 38th International Colloquium on Automata, Languages and Programming, pages 666–677, 2011.

12. N. Kamiyama. Popular Matchings with Two-Sided Preference Lists and Matroid Constraints. Technical Report MI 2016-13, 2016.
13. T. Kavitha. A Size-Popularity Tradeoff in the Stable Marriage Problem. SIAM Journal on Computing, 43(1):52–71, 2014.
14. Z. Király. Better and Simpler Approximation Algorithms for the Stable Marriage Problem. Algorithmica, 60(1):3–20, 2011.

A Appendix

A.1 Reduction from a PCSM$^+$ instance to an SPA instance

An instance of SPA [2] consists of students, projects and lecturers. Each lecturer has an upper bound on the maximum number of students that he/she is willing to advise. Each project has an upper bound on the number of students it can accommodate. Each project is owned by exactly one lecturer. Each student has a preference ordering over a subset of the projects, and each lecturer has a preference over the students.

We detail on the reduction from PCSM$^+$ instance to an SPA instance here. For a resident r in the PCSM$^+$ instance, a corresponding student s_r is introduced in the SPA instance. For each hospital h, a lecturer l_h with capacity $q(h)$ is added in the SPA instance. For each class C_j^h in the classification provided by a hospital h, a project p_j is associated with the lecturer l_h, and the upper-bound of p_j is equal to $q(C_j^h)$. The preference list of l_h is obtained from its corresponding hospital h. If the resident r is the k-th most preferred resident in the preference list of h, then the student s_r is the k-th most preferred student in the preference list of l_h. Similarly, the preference list of a student s_r is created from its corresponding resident r. Let C_j^h be the class that the resident r appears in the classification provided by the k-th most preferred hospital in its preference list, then p_j is the k-th most preferred project in the preference list for s_r. As the classifications associated with every hospital in the PCSM$^+$ instance are a partition over its preference list, there is no ambiguity in describing the preference of the students.