On the (Parameterized) Complexity of Almost Stable Marriage

Sushmita Gupta
The Institute of Mathematical Sciences, HBNI, Chennai, India
sushmitagupta@imsc.res.in

Pallavi Jain
Indian Institute of Technology Jodhpur, India
pallavi@iitj.ac.in

Sanjukta Roy
The Institute of Mathematical Sciences, HBNI, Chennai, India
sanjukta@imsc.res.in

Saket Saurabh
The Institute of Mathematical Sciences, HBNI, Chennai, India
University of Bergen, Norway
saket@imsc.res.in

Meirav Zehavi
Ben Gurion University of the Negev, Beer Sheva, Israel
meiravze@bgu.ac.il

Abstract

In the Stable Marriage problem, when the preference lists are complete, all agents of the smaller side can be matched. However, this need not be true when preference lists are incomplete. In most real-life situations, where agents participate in the matching market voluntarily and submit their preferences, it is natural to assume that each agent wants to be matched to someone in his/her preference list as opposed to being unmatched. In light of the Rural Hospital Theorem, we have to relax the "no blocking pair" condition for stable matchings in order to match more agents. In this paper, we study the question of matching more agents with fewest possible blocking edges. In particular, the goal is to find a matching whose size exceeds that of a stable matching in the graph by at least t and has at most k blocking edges. We study this question in the realm of parameterized complexity with respect to several natural parameters, k, t, d, where d is the maximum length of a preference list. Unfortunately, the problem remains intractable even for the combined parameter $k + t + d$. Thus, we extend our study to the local search variant of this problem, in which we search for a matching that not only fulfills each of the above conditions but is "closest", in terms of its symmetric difference to the given stable matching, and obtain an FPT algorithm.

2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms

Keywords and phrases Stable Matching, Parameterized Complexity, Local Search

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2020.24

Related Version A full version of the paper is available at `https://arxiv.org/pdf/2005.08150.pdf`.

Funding *Sushmita Gupta*: was supported by SERB-Starting Research Grant (SRG/2019/001870).
Saket Saurabh: Received funding from European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant no. 819416), and Swarnajayanti Fellowship grant DST/SJF/MSA-01/2017-18.
Meirav Zehavi: was supported by ISF grant (no.1176/18) and BSF grant no. 2018302).

1 Introduction

Matching various entities to available resources is of great practical importance, exemplified in matching college applicants to college seats, medical residents to hospitals, preschoolers to kindergartens, unemployed workers to jobs, organ donors to recipients, and so on [2, 14, 19, 21]. It is noteworthy that in the applications mentioned above, it is not enough to merely match an entity to any of the available resources. It is imperative, in fact, mission-critical, to create matches that fulfil some predefined notions of compatibility, suitability, acceptability, and so on. Gale and Shapley introduced the fundamental theoretical framework to study such two-sided matching markets in the 1960s. They envisioned a matching outcome as a *marriage* between the members of the two sides, and a desirable outcome representing a *stable marriage*. The algorithm proffered by them has since attained wide-scale recognition as the Gale-Shapley stable marriage/matching algorithm [14]. Stability is one of the acceptability criteria for matching in which an unmatched pair of agent should not prefer each other over their matched partner.

Of the many characteristic features of the two-sided matching markets, there are certain aspects that stand out and are supported by both theoretical and empirical evidence – particularly notable is the curious aspect that for a given market with strict preferences on both sides,[1] no matter what the stable matching outcome is, the specific *number* of resources matched on either side always remains the same. This fact encapsulated by The Rural Hospital's Theorem [30, 31] states that no matter what stable matching algorithm is deployed, the exact *set* (rather than only the number) of resources that are matched on either side is the same. In other words, *there is a trade-off between size and stability such that any increase in size must be paid for by sacrificing stability.* Indeed, it is not hard to find instances in which as much as half of the available resources are unmatched in every stable matching. Such gross underutilization of critical and potentially expensive resources has not gone unaddressed by researchers. In light of The Rural Hospital Theorem, many variations have been considered, some important ones being: enforcing lower and upper capacities, forcing some matches, forbidding some matches, relaxing the notion of stability, and finally foregoing stability altogether in favor of size [2, 3, 7, 16, 22, 34].

We formalize the trade-off mentioned above between size and stability in terms of the Almost Stable Marriage problem. The classical Stable Marriage problem takes as an instance a bipartite graph $G = (A \cup B, E)$, where A and B denote the set of vertices representing the agents on the two sides and E denotes the set of edges representing acceptable matches between vertices on different sides, and a preference list of every vertex in G over its neighbors. Thus, the length of the preference list of a vertex is the same as its degree in the graph. A *matching* is defined as a subset of the set of edges E such that no vertex appears in more than one edge in the matching. An edge in a matching represents a match such that the endpoints of a matching edge are said to be the *matching partners* of each other, and an unmatched vertex is deemed to be self-matched. A matching μ is said to be *stable* in G if there does not exist a *blocking edge* with respect to μ, defined to be an edge $e \in E \setminus \mu$ whose endpoints rank each other higher (in their respective preference lists) than their matching partners in μ.[2] The goal of the Stable Marriage problem is to find a stable matching. We define the Almost Stable Marriage problem as follows.

[1] In most real-life applications, it is unreasonable if not unrealistic to expect each of the agents to rank all the agents on the other side. That is, the graph G is highly unlikely to be complete.
[2] Every candidate is assumed to prefer being matched to any of its neighbors to being self-matched.

Almost Stable Marriage (ASM)
Input: A bipartite graph $G = (A \cup B, E)$, a set $\mathcal{L}$ containing the preference list of each vertex, and non-negative integers k and t.
Question: Does there exist a matching whose size is at least t more than the size of a stable matching in G such that the matching has at most k blocking edges?

In ASM, we hope for a matching that is larger than a stable matching but may contain some blocking edges. The above problem quantifies these two variables: t and k denote the minimum increase in the size and the allowable number of blocking edges, respectively.

We note that Biró et al. [3] considered the problem of finding, among all matchings of maximum size, one that has the fewest blocking edges, and showed the NP-hardness of the problem even when the degree of the graph is at most three. Since one can find a maximum matching and a stable matching in the given graph in polynomial time [27, 14], their NP-hardness result implies NP-hardness for ASM even when the degree is at most three by setting t to be the difference between the size of a maximum matching and the size of a stable matching.

Our Contribution and Methods. We study the parameterized complexity of ASM with respect to parameters k and t; a combination that is not settled by Biró et al. [3]. Our first result exhibits a strong guarantee of intractability. We exhibit parameterized intractability of ASM in a very restrictive setting where the degree of the given graph is three.

▶ **Theorem 1.** *ASM is* W[1]*-hard with respect to $k + t$, even when the maximum degree of the given graph is at most three.*

We prove Theorem 1 by showing a polynomial-time many-to-one parameter preserving reduction from the Multicolored Clique (MCQ, in short) problem to ASM. In the Multicolored Clique problem, given a graph $G = (V, E)$ and a partition of $V(G)$ into k parts, say $V_1, \ldots, V_k$; the goal is to decide the existence of a set $S \subseteq V(G)$ such that $|S \cap V_i| = 1$, for all $i \in [k]$, and $G[S]$ induces a clique, that is, there is an edge between every pair of vertices in $G[S]$. MCQ is known to be W[1]-hard [29, 12] with respect to k.

In light of the intractability result in Theorem 1, we are hard pressed to recalibrate our expectations of what is algorithmically feasible in an efficient manner. Therefore, we consider a local search approach for this problem, in which, instead of finding any matching whose size is at least t larger than the size of stable matching, we also want this matching to be "closest" in terms of its symmetric difference, to a stable matching. Such framework of local search has also been studied for other variants of the Stable Marriage problem by Marx and Schlotter [26, 25]. *We would like to emphasize that the notion of local search used here is different from the classical notion of local search heuristics/algorithms commonly used in practice [33].* We use the notion of local search that is well-defined and widely used in the domain of parameterized complexity, as exemplified by Marx and Schlotter [26, 25], and has also been applied to study several other optimization problems [11, 18, 20, 23, 24, 25, 32]. The question is formally defined as follows.

Local Search-ASM (LS-ASM)
Input: A bipartite graph $G = (A \cup B, E)$, a set $\mathcal{L}$ containing the preference list of every vertex, a stable matching μ, and non-negative integers k, q, and t.
Question: Does there exist a matching η of size at least $|\mu| + t$ with at most k blocking edges such that the symmetric difference between μ and η is at most q?

Unsurprisingly perhaps, the existence of a stable matching in the proximity of which we wish to find a solution does not readily mitigate the computational hardness of the problem, as evidenced by Theorem 2. This result is a consequence of the properties of the reduction used in the proof of Theorem 1. The NP-hardness of LS-ASM also follows from the NP-hardness of ASM as we can set q to be $2n$, the maximum possible size of $\mu \cup \eta$.

▶ **Theorem 2** (♣).[3] *LS-ASM is* W[1]*-hard with respect to* $k + t$*, even when the maximum degree of the given graph is at most three.*

In our quest for a parameterization that makes the problem tractable, we investigate LS-ASM with respect to $k + q + t$.

▶ **Theorem 3** (♣). *LS-ASM is* W[1]*-hard with respect to* $k + q + t$*.*

To prove Theorem 3, we give a polynomial-time many-to-one parameter preserving reduction from MCQ to LS-ASM. In the instance constructed to prove Theorem 1, q is not a function of k. We mimic the idea of gadget construction in that proof and ensure that q is a function of k. However, in this effort, the degree of the graph increases. Consequently, the result in Theorem 3 does not hold for constant degree graphs or even when the degree is a function of k. This trade-off between q and the degree of the graph in the instances that establish intractability is not a coincidence, as implied by our next result.

▶ **Theorem 4.** *There exists an algorithm that, given an instance of LS-ASM, solves the instance in* $2^{\mathcal{O}(q \log d)+o(dq)} n^{\mathcal{O}(1)}$ *time, where* n *is the number of vertices in the given graph, and* d *is the maximum degree of the given graph.*

To prove Theorem 4, we begin by using the technique of *random separation* based on color coding, in which the underlying idea is to highlight the solution that we are looking for with high probability. Suppose that η is a hypothetical solution to the given instance of LS-ASM. Note that to find the matching η, it is enough to find the edges that are in the symmetric difference of μ and η, denoted by $\mu \triangle \eta$. Thus, using the technique of random separation, we wish to highlight the edges in $\mu \triangle \eta$. We achieve this goal using two layers of randomization. The first one separates vertices that appear in $\mu \triangle \eta$, denoted by the set $V(\mu \triangle \eta)$, from its neighbors, by independently coloring vertices 1 or 2. Let the vertices appearing in $V(\mu \triangle \eta)$ be colored 1 and its neighbors that are not in $V(\mu \triangle \eta)$ be colored 2. Observe that the matching partner of the vertices which are not in $V(\mu \triangle \eta)$ is the same in both μ and η. Therefore, we search for a solution locally in vertices that are colored 1. Let G_1 be the graph induced on the vertices that are colored 1. At this stage we use a second layer of randomization on edges of G_1, and independently color each edge with 1 or 2. This separates edges that belong to $\mu \triangle \eta$ (say colored 1) from those that do not belong to $\mu \triangle \eta$. Now for each component of G_1, we look at the edges that have been colored 1, and compute the number of blocking edges, the increase in size and increase in the symmetric difference, if we modify using the μ-alternating paths/cycle that are present in this component. This leads to an instance of the Two-Dimensional Knapsack (2D-KP) problem, which we solve in polynomial time using a known pseudo-polynomial time algorithm for 2D-KP [17]. We derandomize this algorithm using the notion of an *n-p-q-lopsided universal* family [13]. Table 1 summarizes the results for ASM and LS-ASM.

[3] Proofs marked by [♣] are deferred to the full version of the paper.

Table 1 Summary of the results for ASM and LS-ASM. Results in blue row are implied from Theorem 7 in [3].

ASM	LS-ASM
NP-hard for $d = 3$ [3]	NP-hard for $d = 3$ [3]
W[1]-hard for $d = 3$ wrt $k + t$ [Thm. 1]	W[1]-hard wrt $k + q + t$ [Thm. 3] FPT wrt $q + d$ [Thm. 4]

Related Work. We present here some variants of the STABLE MARRIAGE problem which are closely related to our model. In the past, the notion of "almost stability" is defined for the STABLE ROOMMATE problem [1]. In the STABLE ROOMMATE problem, the goal is to find a stable matching in an arbitrary graph. As opposed to STABLE MARRIAGE, in which the graph is a bipartite graph, an instance of STABLE ROOMMATE might not admit a stable matching. Therefore, the notion of almost stability is defined for the STABLE ROOMMATE problem, in which the goal is to find a matching with a minimum number of blocking edges. This problem is known as the ALMOST STABLE ROOMMATE problem. Abraham et al. [1] proved that the ALMOST STABLE ROOMMATE problem is NP-hard. Biro et al. [4] proved that the problem remains NP-hard even for constant-sized preference lists and studied it in the realm of approximation algorithms. Chen et al. [5] studied this problem in the realm of parameterized complexity and showed that the problem is W[1]-hard with respect to the number of blocking edges even when the maximum length of every preference list is five.

Later in 2010, Biró et al. [3] considered the problem of finding, among all matchings of the maximum size, one that has the fewest blocking edges, in a bipartite graph and showed that the problem is NP-hard and not approximable within $n^{1-\epsilon}$, for any $\epsilon > 0$ unless P=NP.

The problem of finding the maximum sized stable matching in the presence of ties and incomplete preference lists, maxSMTI, has striking resemblance with ASM. In maxSMTI, the decision of resolving each tie comes down to deciding who should be at the top of each of tied lists, mirrors the choice we have to make in ASM in rematching the vertices who will be part of a blocking edge in the new matching. Despite this similarity, the W[1]-hardness result presented in [26, Theorem 7] does not yield the hardness result of ASM and LS-ASM as the reduction is not likely to be parameteric in terms of $k + t$ and $k + t + q$, or have the degree bounded by a constant. For other variants of the STABLE MARRIAGE problem, we refer the reader to [6, 21, 15, 19].

2 Preliminaries

Sets. We denote the set of natural numbers $\{1, \ldots, \ell\}$ by $[\ell]$. For two sets X and Y, we use notation $X \triangle Y$ to denote the symmetric difference between X and Y. For any ordered set X, and an appropriately defined value t, $X(t)$ denotes the t^{th} element of the set X. Conversely, suppose that x is t^{th} element of the set X, then $\sigma(X, x) = t$.

Graphs. Let G be an undirected graph. We denote an edge between u and v as uv. The *neighborhood* of a vertex v, denoted by $N_G(v)$, is the set of all vertices adjacent to it. Analogously, the *(open) neighborhood* of a subset $S \subseteq V$, denoted by $N_G(S)$, is the set of vertices outside S that are adjacent to some vertex in S. A *component* of G is a maximal subgraph in which any two vertices are connected by a path. Let H be a subgraph of G. For a component C in H, we set $N_G(C) = N_G(V(C))$. The subscript may be omitted if the graph under consideration is clear from the context.

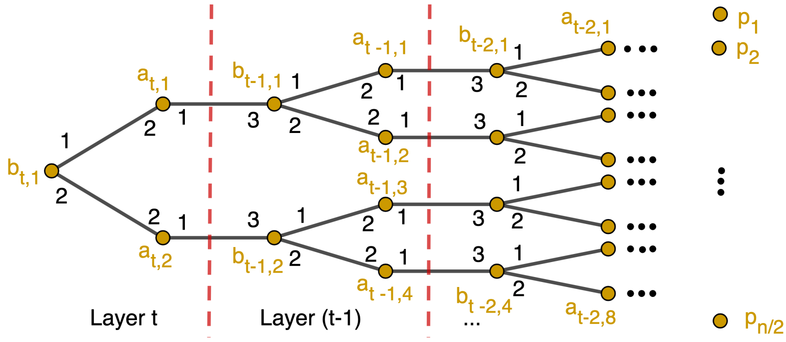

Figure 1 Depiction of the top three layers of special vertices associated with the vertices of G', as explained on page 7, where $t = \log_2(n/2)$. Yellow labels denote vertex labels and numbers on edges denote preferences. Analogously, we can depict the layers of the special vertices associated with the edges of G.

In the preference list of a vertex u, if v appears before w, then we say that u prefers v to w, and denote it as $v \succ_u w$. We call an edge in the graph a *static edge* if its endpoints prefer each other over any other vertex in the graph. For a matching μ, $V(\mu) = \{u, v \colon uv \in \mu\}$. If an edge $uv \in \mu$, then $\mu(u) = v$ and $\mu(v) = u$. A vertex is called *saturated* in a matching μ, if it is an endpoint of one of the edges in the matching μ, otherwise it is an *unsaturated* vertex in μ. If u is an unsaturated vertex in a matching μ, then we write $\mu(u) = \emptyset$. For a matching μ in G, a μ-alternating path (cycle) is a path (cycle) whose edges alternate between matching edges of μ and non-matching edges. A μ-augmenting path is a μ-alternating path that starts and ends at an unmatched vertex in μ.

Unless specified, we will be using all general graph terminologies from the book of Diestel [9]. For parameterized complexity related definitions, we refer the reader to [8, 10, 28].

We conclude this section with a result that is used extensively in our analysis.

▶ **Proposition 1** (♣). *Let μ and μ' denote two matchings in G such that μ is stable and μ' is not. Then, for each blocking edge with respect to μ' we know that at least one of the endpoints has different matching partners in μ and μ'.*

3 W[1]-hardness of ASM

We give a polynomial-time parameter preserving many-to-one reduction from the W[1]-hard problem Multicolored Clique (MCQ) [29, 12] in which we are given a regular graph $G = (V, E)$ and a partition of $V(G)$ into k parts, $V_1, \ldots, V_k$, and the objective is to decide if there exists a subset $S \subseteq V(G)$ such that $|S \cap V_i| = 1$, for each $i \in [k]$, and the induced subgraph $G[S]$ is a clique. Given an instance $\mathcal{I} = (G, (V_1, \ldots, V_k))$ of MCQ, we will next describe the construction of an instance $\mathcal{J} = (G', \mathcal{L}, k', t)$ of ASM.

Construction. We begin by introducing some notations. For any $\{i, j\} \subseteq [k]$, such that $i < j$, we use E_{ij} to denote the set of edges between sets V_i and V_j. For each $i \in [k]$, we have $|V_i| = n = 2^p$, and for each $\{i, j\} \subseteq [k]$, we have $|E_{ij}| = m = 2^{p'}$, for some positive integers p and p' greater than one.[4] We assume that sets V_i (for each $i \in [k]$) and E_{ij} (for

[4] Let m' be the maximum number of edges in any E_{ij}, where $\{i, j\} \subseteq [k]$. Let p' be the smallest positive integer greater than one such that $m' \leq 2^{p'}$. Then, for every $\{i, j\} \subseteq [k]$, add $2^{p'} - |E_{ij}|$ isolated edges

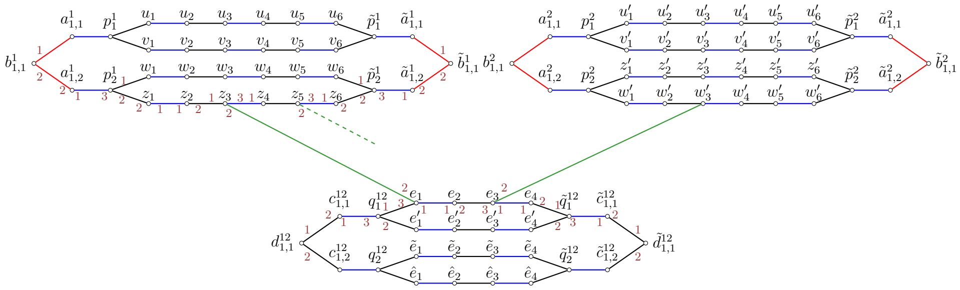

Figure 2 An illustration of construction of graph G' in the proof of W[1]-hardness of ASM for constant sized preference list. Here, blue colored edges belong to the stable matching μ. Here, $n = 4$, $m = 4$, and $r_u = 2$, for all $u \in V(G)$.

each $\{i, j\} \subseteq [k], i < j$) have a canonical order, and thus for an appropriately defined value t, $V_i(t)$ ($E_{ij}(t)$) and $\sigma(V_i, v)$ ($\sigma(E_{ij}, e)$), where $v \in V_i$ and $e \in E_{ij}$, are uniquely defined. For each vertex $u \in V(G)$, let r_u denotes the degree of u in the graph G.

For each $j \in [\log_2(n/2)]$, let $\beta_j = n/2^j$, and $\gamma_j = \beta_j/2$. For each $j \in [\log_2(m/2)]$, let $\rho_j = m/2^j$, and $\tau_j = \rho_j/2$. Next, we are ready to describe the construction of the graph G'.

Base vertices.

- For each vertex $u \in V(G)$, we add $2r_u + 2$ vertices in G', denoted by $\{u_i : i \in [2r_u + 2]\}$, connected via a path: $(u_1, \ldots, u_{2r_u+2})$.
- For each edge $e \in E(G)$, we have four vertices in G', denoted by $\{e_i : i \in [4]\}$, connected via a path: (e_1, e_2, e_3, e_4).

For each vertex $u \in V(G)$, we define a set $\mathscr{E}_u \subseteq V(G')$ as follows. Let $u \in V_i$, for some $i \in [k]$. Then, for any edge $e(= uv) \in E_{ij}$, where $j \in [k]$, $j > i$, we have that the vertex $e_1 \in \mathscr{E}_u$; and for any edge $e(= uv) \in E_{ji}$, where $j \in [k]$, $j < i$, we have that the vertex $e_3 \in \mathscr{E}_u$. Formally,

$$\mathscr{E}_u = \{e_1 \in V(G') : e = uv \in E_{ij}\} \cup \{e_3 \in V(G') : e = uv \in E_{ji}\}$$

We assume that the set $\mathscr{E}_u$ has a canonical ordering. We encode the vertex-edge incidence relation in the graph G as follows: For each vertex $u \in V(G)$ and value $h \in [r_u]$, the vertex u_{2h+1} in G' is a neighbor of the vertex $\mathscr{E}_u(h)$. Thus, the fact that the edge e is incident to a vertex u in G, is captured by the fact that a "copy" of e (namely e_1 or e_3) is adjacent to a "copy" of u in G'.

Special vertices. For each $i \in [k]$, we create the following special vertices associated with the vertices in V_i.

- For each $\ell \in [\beta_1]$, we add vertices p^i_ℓ and $\tilde{p}^i_\ell$ in $V(G')$. Let u and v denote the $2\ell - 1^{st}$ and the $2\ell^{th}$ vertices in V_i, respectively. Then, the vertex p^i_ℓ is a neighbor of vertices u_1 and v_1; and the vertex $\tilde{p}^i_\ell$ is a neighbor of vertices u_{2r_u+2} and v_{2r_v+2} in G'.

(an edge whose endpoints are of degree exactly one) to E_{ij}. Similarly, let n' be the maximum number of vertices in any V_i, where $i \in [k]$. Let p be the smallest positive integer greater than one such that $n' \leq 2^p$. Then, for every $i \in [k]$, add $2^p - |V_i|$ isolated vertices to V_i. Note that if $(G, (V_1, \ldots, V_k))$ was a W[1]-hard instance of MCQ earlier, then so is the modified instance.

- For each $j \in [\log_2(n/2)]$, we add vertices in G' in *layers*, where the value of j gives the layer. Vertices in layer j are $\{b^i_{j,\ell} : \ell \in [\beta_j/2]\} \cup \{a^i_{j,\ell} : \ell \in [\beta_j]\}$. In the 1^{st} layer, $a^i_{1,\ell}$ is a neighbor of p^i_ℓ. In the top layer, i.e., $j = \log_2(n/2)$, $b^i_{j,1}$ is a neighbor of $a^i_{j,1}$ and $a^i_{j,2}$. In intermediate layers, i.e., $1 < j < \log_2(n/2)$, vertex $b^i_{j,\ell}$ is adjacent to two vertices in its layer, namely $a^i_{j,2\ell-1}$, $a^i_{j,2\ell}$ as well as one vertex from layer $j+1$, namely $a^i_{j+1,\ell}$. Refer to Figure (1) for a depiction of two layers.
  Symmetrically, we define vertices $\{\tilde{b}^i_{j,\ell} : \ell \in [\beta_j/2]\} \cup \{\tilde{a}^i_{j,\ell} : \ell \in [\beta_j]\}$ and define similar adjacencies for them as well; details are in Table 2.

For each $\{i,j\} \subseteq [k]$, where $i < j$, we create the following special vertices associated with the edges in E_{ij}.

- For each $\ell \in [\rho_1]$, we add vertices q^{ij}_ℓ and $\tilde{q}^{ij}_\ell$ to $V(G')$.
  Moreover, let e and e' denote the $2\ell - 1^{st}$ and $2\ell^{th}$ elements of E_{ij}, respectively. Then, q^{ij}_ℓ is a neighbor of e_1 and e'_1; and symmetrically $\tilde{q}^{ij}_\ell$ is a neighbor of e_4 and e'_4 in G'.
- As before, for each $h \in [\log_2(m/2)]$, we add vertices in G' in layers, where the value of h indicates the layer. Vertices in layer h are $\{c^{ij}_{h,\ell} : \ell \in [\rho_j/2]\} \cup \{d^{ij}_{h,\ell} : \ell \in [\rho_j]\}$. In the 1^{st} layer, vertex $c^{ij}_{1,\ell}$ is a neighbor of q^{ij}_ℓ. In the top layer, i.e., $h = \log_2(m/2)$, vertex $d^{ij}_{h,1}$ is a neighbor of $c^{ij}_{h,1}$ and $c^{ij}_{h,2}$. In intermediate layers, i.e., $1 < h < \log_2(m/2)$, vertex $d^{ij}_{h,\ell}$ is adjacent to two vertices in its layer, namely $c^{ij}_{h,2\ell-1}$, $c^{ij}_{h,2\ell}$ as well as one vertex from layer $h+1$, namely $c^{ij}_{h+1,\ell}$.
  Symmetrically, we define vertices $\{\tilde{c}^{ij}_{h,\ell} : \ell \in [\rho_j/2]\} \cup \{\tilde{d}^i_{h,\ell} : \ell \in [\rho_j]\}$ and define similar adjacencies for these vertices; details are in Table 2.

Figure 2 illustrates the construction of G'. The preference list of each vertex in G' is presented in Table 2.

Parameter: We set $k' = k^2$, and $t = k + k(k-1)/2$. This completes the construction of an instance of ASM. Clearly, this construction can be carried out in polynomial time. Since $|V_i| = n$ and $|E_{ij}| = m$, for every $\{i,j\} \subseteq [k]$, we have $2nk + 4mk(k-1)$ many base vertices and $4nk + 2mk(k-1) - 3k - 3k^2$ many special vertices. Thus, in total we have

$$|V(G')| = 6mk(k-1) + 6nk - 3k - 3k^2. \tag{I}$$

The rest of the proof of Theorem 1 is deferred to the full version of the paper.

4 FPT Algorithm for LS-ASM

In this section, we give an FPT algorithm for LS-ASM with respect to $q + d$ (Theorem 4). Recall that d is the degree of the graph G, and q is the symmetric difference between a solution matching and the given stable matching μ. Before presenting our algorithm, we prove that there exists a solution, γ, to $(G, \mathcal{L}, \mu, k, q, t)$ such that in every component of $G[V(\mu \triangle \gamma)]$, the number of γ-edges (edges that are in γ) is more than the number of μ-edges in this component. We will need such a solution for a technical purpose which will be cleared later in Phase III of the algorithm.

▶ **Lemma 5.** *There exists a solution γ to $(G, \mathcal{L}, \mu, k, q, t)$ such that for every component C of $G[V(\mu \triangle \gamma)]$, $|E(C) \cap \gamma| > |E(C) \cap \mu|$.*

The proof of Lemma 5 follows by starting with a solution γ and then replacing the edges in μ with the edges in γ only in those components of $G[V(\mu \triangle \gamma)]$, where $|\gamma| > |\mu|$.

We begin with the description of a randomized algorithm which will be derandomized later using *n-p-q-lopsided universal* family [13]. Our algorithm has three phases: Vertex Separation, Edge Separation, and Size-Fitting. Given an instance $(G, \mathcal{L}, \mu, k, q, t)$ of LS-ASM, we proceed as follows.

Phase I: Vertex Separation. We start with the following assumption.

Table 2 Preference lists in the proof of Theorem 1. Here, for a set S, the notation $\langle S \rangle$ denotes the order of preference over the vertices in this set.

For each vertex $u \in V_i$, where $i \in [k]$, we have the following preferences:		
u_1:	$\langle u_2, p^i_{\lceil \ell/2 \rceil} \rangle$	where for some $\ell \in [n]$, $u = V_i(\ell)$.
u_{2h+1}:	$\langle u_{2h}, \mathscr{E}_u(h), u_{2h+2} \rangle$	where $h \in [r_u]$
u_{2h}:	$\langle u_{2h-1}, u_{2h+1} \rangle$	where $h \in [r_u]$
u_{2r_u+2}:	$\langle u_{2r_u+1}, \tilde{p}^i_{\lceil \ell/2 \rceil} \rangle$	where for some $\ell \in [n]$, $u = V_i(\ell)$.
For the special vertices associated with V_i, we have the following preferences:		
p^i_ℓ:	$\langle u_1, v_1, a^i_{1,\ell} \rangle$	where $\ell \in [n/2]$, $u = V_i(2\ell - 1)$ and $v = V_i(2\ell)$
$\tilde{p}^i_\ell$:	$\langle u_{2r_u+2}, v_{2r_v+2}, \tilde{a}^i_{1,\ell} \rangle$	where $\ell \in [n/2]$, $u = V_i(2\ell - 1)$ and $v = V_i(2\ell)$
$a^i_{1,\ell}$:	$\langle p^i_\ell, b^i_{1,\lceil \ell/2 \rceil} \rangle$	where $\ell \in [n/2]$
$\tilde{a}^i_{1,\ell}$:	$\langle \tilde{p}^i_\ell, \tilde{b}^i_{1,\lceil \ell/2 \rceil} \rangle$	where $\ell \in [n/2]$
$a^i_{j,\ell}$:	$\langle b^i_{j-1,\ell}, b^i_{j,\lceil \ell/2 \rceil} \rangle$	where $j \in [\log_2(n/2)] \setminus \{1\}$ and $\ell \in [n/2^j]$
$\tilde{a}^i_{j,\ell}$:	$\langle \tilde{b}^i_{j-1,\ell}, \tilde{b}^i_{j,\lceil \ell/2 \rceil} \rangle$	where $j \in [\log_2(n/2)] \setminus \{1\}$ and $\ell \in [n/2^j]$
$b^i_{j,\ell}$:	$\langle a^i_{j,2\ell-1}, a^i_{j,2\ell}, a^i_{j+1,\ell} \rangle$	where $j \in [\log_2(n/2) - 1]$ and $\ell \in [n/2^{j+1}]$
$\tilde{b}^i_{j,\ell}$:	$\langle \tilde{a}^i_{j,2\ell-1}, \tilde{a}^i_{j,2\ell}, \tilde{a}^i_{j+1,\ell} \rangle$	where $j \in [\log_2(n/2) - 1]$ and $\ell \in [n/2^{j+1}]$
$b^i_{j,1}$:	$\langle a^i_{j,1}, a^i_{j,2} \rangle$	where $j = \log_2(n/2)$
$\tilde{b}^i_{j,1}$:	$\langle \tilde{a}^i_{j,1}, \tilde{a}^i_{j,2} \rangle$	where $j = \log_2(n/2)$
For each edge $e \in E_{ij}$, $1 \leq i < j \leq k$, we have the following preferences:		
e_1:	$\langle e_2, u_{2h+1}, q^{ij}_{\lceil \ell/2 \rceil} \rangle$	where for some $\ell \in [m]$, $e = uv = E_{ij}(\ell)$ s.t. $u \in V_i$ and for some $h \in [r_u]$, $e_1 = \mathscr{E}_u(h)$
e_2:	$\langle e_1, e_3 \rangle$	
e_3:	$\langle e_4, v_{2h+1}, e_2 \rangle$	where $e = uv$ s.t $v \in V_j$ and for some $h \in [r_v]$, $e_3 = \mathscr{E}_v(h)$.
e_4:	$\langle e_3, \tilde{q}^{ij}_{\lceil \ell/2 \rceil} \rangle$	where for some $\ell \in [m]$, $e = uv = E_{ij}(\ell)$
For the special vertices associated with E_{ij}, we have the following preferences:		
q^{ij}_ℓ:	$\langle e_1, e'_1, c^{ij}_{1,\ell} \rangle$	where $\ell \in [m/2]$, $e = \sigma(E_{ij}, 2\ell - 1)$ and $e' = \sigma(E_{ij}, 2\ell)$
$\tilde{q}^{ij}_\ell$:	$\langle e_4, e'_4, \tilde{c}^{ij}_{1,\ell} \rangle$	where $\ell \in [m/2]$, $e = \sigma(E_{ij}, 2\ell - 1)$ and $e' = \sigma(E_{ij}, 2\ell)$
$c^{ij}_{1,\ell}$:	$\langle q^{ij}_\ell, d^{ij}_{1,\lceil \ell/2 \rceil} \rangle$	where $\ell \in [m/2]$
$\tilde{c}^{ij}_{1,\ell}$:	$\langle \tilde{q}^{ij}_\ell, \tilde{d}^{ij}_{1,\lceil \ell/2 \rceil} \rangle$	where $\ell \in [m/2]$
$c^{ij}_{h,\ell}$:	$\langle d^{ij}_{h-1,\ell}, d^{ij}_{h,\lceil \ell/2 \rceil} \rangle$	where $h \in [\log_2(m/2)] \setminus \{1\}, \ell \in [m/2^h]$
$\tilde{c}^{ij}_{h,\ell}$:	$\langle \tilde{d}^{ij}_{h-1,\ell}, \tilde{d}^{ij}_{h,\lceil \ell/2 \rceil} \rangle$	where $h \in [\log_2(m/2)] \setminus \{1\}$ and $\ell \in [m/2^h]$
$d^{ij}_{h,\ell}$:	$\langle c^{ij}_{h,2\ell-1}, c^{ij}_{h,2\ell}, c^{ij}_{h+1,\ell} \rangle$	where $h \in [\log_2(m/2) - 1]$ and $\ell \in [m/2^{h+1}]$
$\tilde{d}^{ij}_{h,\ell}$:	$\langle \tilde{c}^{ij}_{h,2\ell-1}, \tilde{c}^{ij}_{h,2\ell}, \tilde{c}^{ij}_{h+1,\ell} \rangle$	where $h \in [\log_2(m/2) - 1]$ and $\ell \in [m/2^{h+1}]$
$d^{ij}_{h,1}$:	$\langle c^{ij}_{h,1}, c^{ij}_{h,2} \rangle$	where $h = \log_2(m/2)$
$\tilde{d}^{ij}_{h,1}$:	$\langle \tilde{c}^{ij}_{h,1}, \tilde{c}^{ij}_{h,2} \rangle$	where $h = \log_2(m/2)$

Throughout this section we assume that there exists solution η, and everything will be defined with respect to η. In fact, we assume that η is a *hypothetical* solution to $(G, \mathcal{L}, \mu, k, q, t)$ such that in every component of $G[V(\mu \triangle \eta)]$, the number of η-edges is more than the number of μ-edges in this component, that is, η satisfies the property specified in Lemma 5.

We start by defining a notion of good coloring.

▶ **Definition 6.** *A function $f : V(G) \to \{0, 1\}$ is called a* good coloring*, if the following properties are satisfied.*

1. *Every vertex in $V(\mu \triangle \eta)$ is colored 1.*
2. *Let* border *be the set of neighbors of the vertices in $V(\mu \triangle \eta)$ outside the set $V(\mu \triangle \eta)$, that is,* border $= N_G(V(\mu \triangle \eta))$, *and* bordermates *be the set of matching partners (if they exist) of the vertices in* border *in μ. Every vertex in* border $\cup$ bordermates *is colored 2.*

We will show that a random function f that assigns each vertex of the graph G independently with color 1 or 2 with probability[5] 1/2 each is a good function with probability depending only on q and d. In particular, we can say the following about f.

- Every vertex in $V(\mu \triangle \eta)$ is colored 1 w.p. at least $\frac{1}{2^{2q}}$.
- Every vertex in border $\cup$ bordermates is colored 2 w.p. at least $\frac{1}{2^{4qd}}$. To see this, note that $|\mu \triangle \eta| \leq q$ and the maximum degree of a vertex in the graph G is d, and so $|\text{border} \cup \text{bordermates}| \leq 2|\text{border}| = 2|N_G(V(\mu \triangle \eta))| \leq 4qd$.

For each $i \in [2]$, let V_i denotes the set of vertices of the graph G that are colored i using the function f. Summarizing the above mentioned properties we get the following.

▶ **Lemma 7.** *Let V_1, V_2,* border *and* bordermates *be as defined above. Then, w.p. at least $\frac{1}{2^{2q+4qd}}$, $V(\mu \triangle \eta) \subseteq V_1$ and* border $\cup$ bordermates $\subseteq V_2$*. Thus, f is a good coloring w.p. at least $\frac{1}{2^{2q+4qd}}$.*

Due to Lemma 7, we have the following:

▶ **Corollary 8.** *Every component in $G[V(\mu \triangle \eta)]$ is a component in $G[V_1]$ w.p. at least $\frac{1}{2^{2q+4qd}}$.*

The proof of Corollary 8 follows from the fact that $V(\mu \triangle \eta) \subseteq V_1$ and border $= N_G(V(\mu \triangle \eta))$ is a subset of V_2 w.p. at least $\frac{1}{2^{2q+4qd}}$. Due to Corollary 8, if there exists a component C in $G[V_1]$ containing a vertex $u \in V(G)$ that is saturated in μ, such that $\mu(u) \notin C$, then C is not a component in $G[V(\mu \triangle \eta)]$. This leads to the following definition. A component C in $G[V_1]$ is called a *colored-component*, if for every vertex $v \in C$, we have that $\mu(v) \in C$. Thus, we get the following lemma.

▶ **Lemma 9.** *Let G be a graph and $f : V(G) \to \{0, 1\}$ be a good function. Then, every component C of $G[V(\mu \triangle \eta)]$ is also a component of $G[V_1]$ and further it is a colored-component.*

Let (G, f) be a pair such that G is the input graph and f is a good coloring function on $V(G)$. We call such (G, f) as a *colored instance.*

In light of Corollary 8, to find $\mu \triangle \eta$, in Phase II, we color the edges of $G[V_1]$ in order to identify the components of the graph that *only contain edges of $\mu \triangle \eta$*. Let $G_1 = G[V_1]$ and $G' = G_1[V(\mu \triangle \eta)]$.

[5] Henceforth, we will use the shortened form w.p. for “with probability”.

Phase II: Edge Separation. We first define a notion of *edge-colored instance.*

▶ **Definition 10.** *Let $f : V(G) \to \{0,1\}$ and $g : E(G) \to \{$Red, Green, Blue$\}$ be two functions. An instance (G, f, g) is called an* edge-colored instance *if the following properties are satisfied.*
1. *(G, f) is a colored instance.*
2. *Every edge in $\mu \triangle \eta$ is colored Red.*
3. *Every edge in $E(G') \setminus (\mu \triangle \eta)$ is colored Green.*
4. *Every edge in $E(G) \setminus E(G_1)$ is colored Blue.*

Given a colored instance (G, f), we select a function g, as explained below, such that (G, f, g) becomes an edge-colored instance with high probability.

Let g be a function that colors each edge of the subgraph G_1 independently with colors Red or Green with probability $1/2$ each. Furthermore, g colors every edge in $E(G) \setminus E(G_1)$ with Blue.

The following properties hold for the graph G_1 that is colored using the function g:

- Every edge in $\mu \triangle \eta$ is colored Red with probability at least $\frac{1}{2^q}$.
- Every edge in $E(G') \setminus (\mu \triangle \eta)$ is colored Green with probability at least $\frac{1}{2^{2qd}}$, because $|V(\mu \triangle \eta)| \leq 2q$ and d is the maximum degree of a vertex in the graph G, so $|E(G')| \leq 2qd$.
- Every edge in $E(G) \setminus E(G_1)$ has been colored Blue w.p. 1.

For $i \in \{$Red, Green, Blue$\}$, let E_i denotes the set of edges of the graph G that are colored i using the function g. Then, due to the above mentioned coloring properties of the graph G_1, we have the following result.

▶ **Lemma 11.** *Let (G, f) be a colored instance. Furthermore, let G', E_{Red}, E_{Green}, and E_{Blue} be as defined above. Then, w.p. at least $\frac{1}{2^{q+2qd}}$, $\mu \triangle \eta \subseteq E_{\mathsf{Red}}$, $E(G') \setminus (\mu \triangle \eta) \subseteq E_{\mathsf{Green}}$, and $E(G) \setminus E(G_1) \subseteq E_{\mathsf{Blue}}$. Thus, (G, f, g) is an edge-colored instance w.p. at least $\frac{1}{2^{q+2qd}}$.*

Note that the edges in $\mu \triangle \eta$ form vertex-disjoint maximal μ-alternating paths/cycles. A component may have *several* μ-alternating paths and cycles. Let C be a colored-component. In what follows, we provide conditions such that if C satisfies either of them, then they *do not belong* to $G[V(\mu \triangle \eta)]$. Such a colored-component is called *malformed.*
1. If the set of Red edges in C do not form vertex disjoint maximal μ-alternating paths or cycles, then the component does not belong to $G[V(\mu \triangle \eta)]$.
2. Furthermore, due to our assumption on the hypothetical solution η, if the number of Red edges in C that are not in μ is at most the number of Red edges in C that are in μ, then C does not belong to $G[V(\mu \triangle \eta)]$.
3. If C does not have any Red edge, then it does not belong to $G[V(\mu \triangle \eta)]$.

A component C in G_1 that is not malformed is called an *edge-colored-component* (edge-colored-comp).

The next observation follows from the properties of an edge-colored component.

▶ **Observation 1.** *Let (G, f, g) be an edge-colored instance. Then, for every* edge-colored-comp *C of G_1, the following holds: (a) The set of Red colored edges form a collection of μ-alternating path/cycle; and (b) every vertex in C is incident to at least one Red edge.*

Let $\mathscr{C}_{\mathsf{ecc}}$ be the set of components of G_1 that are edge-colored-comp. In light of Observation 1, our goal is reduced to finding a family of components, $\mathscr{C}$, in $\mathscr{C}_{\mathsf{ecc}}$ that contain the edges of $\mu \triangle \eta$. Indeed, to obtain a matching of size at least $|\mu| + t$, we need to choose $t' \leq t$ components of $G[V_1]$ that have μ-augmenting paths (a μ-alternating path starting and ending with edges not in μ). However, choosing t' components arbitrarily might lead to a large number of blocking edges in the solution matching. Thus, to choose the components appropriately, we move to Phase III. In particular we show that if (G, f, g) is an edge-colored instance, then we can solve the problem in polynomial time.

Phase III: Size-Fitting with respect to g. Let $(G, \mathcal{L}, \mu, k, q, t)$ be an instance to LS-ASM and η be a hypothetical solution to the problem that satisfies the condition in Lemma 5. Further, let (G, f, g) be an edge-colored instance and $\mathscr{C}_{\mathsf{ecc}}$ be the set of components of G_1 that are edge-colored-comp.

We reduce our problem to Two-Dimensional Knapsack (2D-KP), and after that use an algorithm for 2D-KP, described in Proposition 2, as a subroutine.

> Two-Dimensional Knapsack (2D-KP)
> **Input:** A set of tuples, $\mathcal{X} = \{(a_i, b_i, p_i) \in \mathbb{N}^3 : i \in [n]\}$, and non-negative integers c_1, c_2 and p.
> **Question:** Does there exist a set $Z \subseteq [n]$ such that $\sum_{i \in Z} a_i \leq c_1$, $\sum_{i \in Z} b_i \leq c_2$, and $\sum_{i \in Z} p_i \geq p$?

▶ **Proposition 2.** [17] *There exists an algorithm $\mathcal{A}$ that given an instance $(\mathcal{X}, c_1, c_2, p)$ of 2D-KP, in time $\mathcal{O}(nc_1c_2)$, outputs a solution if it is a* Yes*-instance of 2D-KP; otherwise $\mathcal{A}$ outputs* "no"*.*

Construction 2D-Knapsack. We construct an instance of 2D-KP as follows. Let $C_1, \ldots, C_\ell$ be the components in $\mathscr{C}_{\mathsf{ecc}}$. Intuitively, we construct a family of tuples $\mathcal{X} = \{(k_i, q_i, t_i) : i \in [\ell]\}$ such that k_i denotes the number of blocking edges that we encounter if we add edges that are not in μ but are present in μ-alternating paths/cycles in C_i to our solution. Similarly, q_i and t_i denote the number of edges in the symmetric difference and the increase in the size of the matching due to this alternation operation. By our choice of the components in $\mathscr{C}_{\mathsf{ecc}}$ all these values are positive integers. Indeed, this is why we selected a hypothetical solution with an additional property. Next, we describe the construction of an instance of 2D-KP.

For each $i \in [\ell]$, let q_i be the number of Red colored edges in C_i and $t_i = q_i - 2|\mu_i|$ where μ_i denotes the edges of μ in C_i. Next, to compute k_i, for each $i \in [\ell]$, we construct a matching ξ_i as follows. We add all the Red colored edges in C_i that are not in μ to ξ_i. Next, we make another matching Γ_i, that has all the edges in ξ_i, and additionally, we add all the edges in μ to Γ_i whose both endpoints are outside the components in $\mathscr{C}_{\mathsf{ecc}}$, and at least one of the endpoints is a neighbor of a vertex in C_i. Clearly, Γ_i is a matching in the graph G. To ease notation, we let G^i denote the graph $G[V(\Gamma_i) \cup V(C_i) \cup N_G(V(C_i))]$. We set k_i as the number of blocking edges with respect to Γ_i in the graph G^i. Basically, the graph G^i contains all the vertices in C_i, their neighbors in border, the μ-partners of these border vertices in bordermates, and the neighbors of C_i which are unsaturated in μ. That is, the number of blocking edges (with respect to Γ_i) incident on the vertices in the set $V(\xi_i)$ is k_i in G^i. To see this note that there is *no blocking edge* with both endpoints in $V(\Gamma_i \setminus \xi_i)$

(Proposition 1). The only reason to define Γ_i is to define the value of k_i in a clean fashion. We next state a simple lemma that shows that no blocking edge is counted twice.

▶ **Lemma 12** (Locally Pairwise Disjoint Blocking Edges). *Let $(G, \mathcal{L}, \mu, k, q, t)$ be an instance of LS-ASM and (G, f, g) be an edge-colored instance. Further, let $C_i, C_j \in \mathscr{C}_{\mathsf{ecc}}$, $i \neq j$, and for $\ell \in \{i, j\}$, B_ℓ denote the set of blocking edges with respect to Γ_ℓ in $G[V(\Gamma_\ell) \cup V(C_\ell) \cup N(V(C_\ell))]$. Then, $B_i \cap B_j = \emptyset$.*

Proof. Due to the construction of the matching Γ_i, for all the blocking edges in B_i, at least one of its endpoints is in C_i. Similarly, for all the blocking edges in B_j, at least one of its endpoints is in C_j. Since C_i are C_j are distinct components in $\mathscr{C}_{\mathsf{ecc}}$, we infer $B_i \cap B_j = \emptyset$. ◀

Let $\mathcal{X} = \{(k_i, q_i, t_i) : i \in [\ell]\}$. This completes the construction of an instance $(\mathcal{X}, k, q, t)$ of 2D-KP. We invoke the algorithm $\mathcal{A}$ given in Proposition 2 on the instance $(\mathcal{X}, k, q, t)$ of 2D-KP. If $\mathcal{A}$ returns a set Z, then we return "yes". Otherwise, we report *failure* of the algorithm. It is relatively straightforward to create the solution $\hat{\eta}$ when the answer is "yes". Next, we prove the correctness of Phase III.

▶ **Lemma 13.** *Let (G, f, g) be an edge-colored instance. Then, $(\mathcal{X}, k, q, t)$ is a yes-instance of 2D-KP.*

Proof. Since (G, f, g) is an edge-colored instance, due to the definition of edge-colored instance (Definition 10), (G, f) is a colored-instance. Thus, due to the definition of a colored-instance and Lemma 9, every component in $G[V(\mu \triangle \eta)]$ is also a component in G_1.

Clearly, for every component C in $G[V(\mu \triangle \eta)]$, if a vertex $u \in C$, then $\mu(u) \in C$. Therefore, all the components in $G[V(\mu \triangle \eta)]$ are colored component. Next, we note that due to Definition 10, all the edges in the set $\mu \triangle \eta$ are colored Red. Thus, for every component C in $G[V(\mu \triangle \eta)]$, Red colored edges in C form vertex disjoint maximal μ-alternating paths/cycles. Further, every component C in $G[V(\mu \triangle \eta)]$ has at least one Red edge. Also, due to our choice of η, the number of Red edges in C, which are not in μ, are less than the one that are in μ. Therefore, all the components in $G[V(\mu \triangle \eta)]$ are edge-colored-comp. Without loss of generality, let $C_1, \dots, C_{\hat{\ell}}$ be the components in $\mathscr{C}_{\mathsf{ecc}}$ that are also in $G[V(\mu \triangle \eta)]$. Let us note that $\mathscr{C}_{\mathsf{ecc}}$ may contain *several other components*. Let $S = \{i \in [\hat{\ell}] : (k_i, q_i, t_i) \in \mathcal{X}\}$. We claim that S is a solution to $(\mathcal{X}, k, q, t)$.

Due to the construction of the instance $(\mathcal{X}, k, q, t)$, and the facts that η is a solution to $(G, \mathcal{L}, \mu, k, q, t)$ and (G, f, g) is an edge-colored instance, clearly, $\sum_{i \in S} q_i \leq q$ and $\sum_{i \in S} t_i \geq t$. We next show that $\sum_{i \in S} k_i \leq k$. Recall the definition of ξ_i and Γ_i. We show that every blocking edge with respect to Γ_i in the graph G^i is also a blocking edge with respect to η in the graph G. Let uv be a blocking edge with respect to Γ_i in the graph G^i. Then, $v \succ_u \Gamma_i(u)$ and $u \succ_v \Gamma_i(v)$. Due to Proposition 1 and the definition of the matching Γ_i, at least one of the endpoint of the edge uv is in the component C_i. Without loss of generality, let $u \in V(C_i)$. Since C_i is also a component in $G[V(\mu \triangle \eta)]$, we can infer that $\eta(u) = \Gamma_i(u)$. If the vertex v is also in the component C_i, then using the same argument as above, we know that $\eta(v) = \Gamma_i(v)$. Thus, uv is also a blocking edge with respect η in the graph G. Suppose that $v \notin V(C_i)$. Then, since $v \in V(\Gamma_i) \cup V(C_i)$, $\Gamma_i(v) = \mu(v)$. Since C_i is a component in $G[V(\mu \triangle \eta)]$ and $u \in V(C_i)$ but $v \notin C_i$, we can infer that $\eta(v) = \mu(v)$. Since u and v have same matching partners in both the matchings η and Γ_i, we can infer that uv is also a blocking edge with respect η in the graph G. Since k_i is the the number of blocking edges with respect to Γ_i, we infer $\sum_{i \in S} k_i \leq k$. Hence, $(\mathcal{X}, k, q, t)$ is a YES-instance of 2D-KP. ◀

▶ **Lemma 14.** *Suppose that (χ, k, q, t) is a YES-instance of 2D-KP. Then, $(G, \mathcal{L}, \mu, k, q, t)$ is a YES-instance of LS-ASM.*

Proof. Suppose that the algorithm $\mathcal{A}$ in Proposition 2 returns the set Z. Given the set Z, we obtain the matching $\tilde{\eta}$ as follows. Let $Z(\mathscr{C})$ denotes the family of components in $\mathscr{C}_{\mathsf{ecc}}$ corresponding to the indices in Z. Formally, $Z(\mathscr{C}) = \{C_i \in \mathscr{C}_{\mathsf{ecc}} : i \in Z\}$. For each component $C \in Z(\mathscr{C})$, we add all the Red edges in C that are not in μ, to $\tilde{\eta}$. That is, $\tilde{\eta} = \cup_{i\in Z}\xi_i$. Additionally, we add all the edges in μ to $\tilde{\eta}$ whose both endpoints are outside the components in $Z(\mathscr{C})$. We next prove that $\tilde{\eta}$ is a solution to $(G, \mathcal{L}, \mu, k, q, t)$.

▷ Claim 15. $\tilde{\eta}$ is a matching.

Proof. Towards the contradiction, suppose that $uv, uw \in \tilde{\eta}$, that is, there exists a pair of edges in $\tilde{\eta}$ that shares an endpoint. Due to Observation 1, in every component of $Z(\mathscr{C})$, the Red edges form μ-alternating path/cycle. Therefore, uv and uw both cannot be in a component of $Z(\mathscr{C})$. Suppose that uv is in a component C in $Z(\mathscr{C})$, but uw does not belong to C. Then, due to the construction of $\tilde{\eta}$, $uw \in \mu$. This contradicts Lemma 9, as C is a component in $\mathscr{C}_{\mathsf{ecc}}$. If uv and uw are outside the components in $Z(\mathscr{C})$, then due to the construction of $\tilde{\eta}$, uv and uw both are in μ. This contradicts that μ is a matching. ◁

▷ Claim 16. $|\mu \triangle \tilde{\eta}| \le q$ and $|\tilde{\eta}| \ge |\mu| + t$.

Proof. Let C be a component in $Z(\mathscr{C})$. Let $E_{\mathsf{Red}}(C)$ denote the set of Red edges in the component C. For each component $C_i \in Z(\mathscr{C})$, let $\mu_i = \mu \cap E_{\mathsf{Red}}(C_i)$, that is, μ_i is the set of Red edges in C_i that are in μ. Let $\tilde{\mu}$ be the set of edges in μ that does not belong to any component in $Z(\mathscr{C})$. Thus, $\mu = \uplus_{C_i\in Z(\mathscr{C})}\mu_i \uplus \tilde{\mu}$. Due to the construction of $\tilde{\eta}$, we have that $\tilde{\eta} = \uplus_{C_i\in Z(\mathscr{C})}(E_{\mathsf{Red}}(C_i) \setminus \mu_i) \uplus \tilde{\mu}$. Thus, $\mu \triangle \tilde{\eta} = \uplus_{C_i\in Z(\mathscr{C})} E_{\mathsf{Red}}(C_i)$. Hence,

$$|\mu \triangle \tilde{\eta}| = \sum_{C_i\in Z(\mathscr{C})} |E_{\mathsf{Red}}(C_i)| = \sum_{i\in Z} q_i$$

as $q_i = |E_{\mathsf{Red}}(C_i)|$ for every component $C_i \in \mathscr{C}_{\mathsf{ecc}}$. Since S is a solution to $(\mathcal{X}, k, q, t)$, $\sum_{i\in Z} q_i \le q$. Therefore, $|\mu \triangle \tilde{\eta}| \le q$. Next, we show that $|\tilde{\eta}| \ge |\mu| + t$. Due to the construction of $\tilde{\eta}$, we know that

$$|\tilde{\eta}| = |\tilde{\mu}| + \sum_{C_i\in Z(\mathscr{C})} |E_{\mathsf{Red}}(C_i) \setminus \mu_i| \quad = |\tilde{\mu}| + \sum_{C_i\in Z(\mathscr{C})} (q_i - |\mu_i|) = |\tilde{\mu}| + \sum_{C_i\in Z(\mathscr{C})} (t_i + |\mu_i|)$$

as $t_i = q_i - 2|\mu_i|$. Since $\sum_{i\in Z} t_i \ge t$, we obtained that $|\tilde{\eta}| \ge |\mu| + t$. ◁

▷ Claim 17. There are at most k blocking edges with respect to $\tilde{\eta}$.

Proof. Due to the construction of the matching $\tilde{\eta}$ and Proposition 1, we know that if uv is a blocking edge with respect to $\tilde{\eta}$, then at least one of its endpoint, that is vertex u or v, is in C_i, for some $C_i \in Z(\mathscr{C})$. Without loss of generality, let the vertex u is in the component C_i. Then, due to the definition of the matching Γ_i and the construction of the matching $\tilde{\eta}$, we have that $\tilde{\eta}(u) = \Gamma_i(u)$. Now, if v is also in the component C_i, then using the same argument $\tilde{\eta}(v) = \Gamma_i(v)$. Suppose that v is not in the component C_i, then its μ-partner, that is $\mu(v)$ is also not present in $\mathscr{C}_{\mathsf{ecc}}$ due the the definition of colored-components. Thus, $\tilde{\eta}(v) = \Gamma_i(v) = \mu(v)$. Since the matching partners of u and v are same in both the matchings η and Γ_i, we have uv is also a blocking edge with respect to matching Γ_i. Thus, every blocking edge with respect to $\tilde{\eta}$ is also a blocking edge with respect to Γ_i, for some $C_i \in Z(\mathscr{C})$. Recall that k_i is the number of blocking edges with respect to Γ_i in G^i. Therefore, the number of blocking edges with respect to $\tilde{\eta}$ is at most $\sum_{i\in Z} k_i \le k$. ◁

Due to Claims 15, 16, and 17, we can infer that $\tilde{\eta}$ is a solution to $(G, \mathcal{L}, \mu, k, q, t)$. ◀

Due to Lemmas 7 and 11, we obtain a polynomial-time randomized algorithm for LS-ASM which succeeds with probability $\frac{1}{2^{3q+6qd}}$. Therefore, by repeating the algorithm independently 2^{3q+6dq} times, where n is the number of vertices in the graph, we obtain the following result:

▶ **Theorem 18.** *There exists a randomized algorithm that given an instance of LS-ASM runs in $2^{3q+6dq}n^{\mathcal{O}(1)}$ time, where n is the number of vertices in the given graph, and either reports a failure or outputs "yes". Moreover, if the algorithm is given a* YES*-instance of the problem, then it returns "yes" with a constant probability.*

Proof. Let $(G, \mathcal{L}, \mu, k, q, t)$ be an instance to LS-ASM and η be a hypothetical solution to the problem. If $(G, \mathcal{L}, \mu, k, q, t)$ is a YES-instance then by Lemmas 7 and 11, we can get an edge-colored instance, (G, f, g), w.p. at least $\frac{1}{2^{3q+6qd}}$. Given an edge-colored instance (G, f, g), we apply **Construction 2D-Knapsack** and construct an instance of 2D-KP with a family of tuples $\mathcal{X} = \{(k_i, q_i, t_i) : i \in [\ell]\}$. Here, $(\mathcal{X}, k, q, t)$ is a yes-instance of 2D-KP. We can solve the instance in polynomial time using Proposition 2. Correctness of this step follows from Lemmas 13 and 14. Thus, if $(G, \mathcal{L}, \mu, k, q, t)$ is a YES-instance, then we return that it is a YES-instance with probability at least $\frac{1}{2^{3q+6qd}}$. Indeed, if $(G, \mathcal{L}, \mu, k, q, t)$ is a NO-instance, then we return that it is a NO-instance with probability 1. Thus, to boast the success probability to a constant, we repeat the algorithm independently $2^{3q+6dq}(\log n)^{\mathcal{O}(1)}$ times, where n is the number of vertices in G. Indeed, the success probability is at least

$$1 - \left(1 - \frac{1}{2^{3q+6qd}}\right)^{2^{3q+6dq}(\log n)^{\mathcal{O}(1)}} \geq 1 - \frac{1}{n^{\mathcal{O}(1)}}.$$

This concludes the proof. ◀

The derandomization of the algorithm is in the full version.

5 In Conclusion

In this paper, we initiated the study of the computational complexity of the trade-off between size and stability through the lenses of both multivariate analysis and local search. Since ASM is NP-hard for a graph in which every vertex has degree at most three, the natural question that arises here is: Is ASM polynomial-time solvable for the graph in which every vertex has degree at most two? It is worth mentioning that there is a fairly straightforward dynamic programming algorithm that solves this question in polynomial time. The basis idea is as follows. This graph, quite clearly, is a disjoint union of paths and cycles.

Consider a hypothetical solution η in the path or cycle X_n in G. Suppose that we know the submatching of η, call it η', that is contained in a subpath of X_n as well as the subset of blocking edges with respect to η that are in this subpath. Then, we can extend η' to η by keeping all the necessary partial solutions. We can briefly sketch this idea as follows. Suppose that η_i is a matching in the subpath $P_i = (1, \ldots, i)$. We want to extend η_i for the subpath P_{i+1}. If v_i is saturated in η_i, then we cannot add edge v_iv_{i+1} to η_{i+1} and we can easily check if it is a blocking edge with respect to η_{i+1} in P_{i+1}. If v_i is unsaturated in η_i, then we have two possibilities: edge v_iv_{i+1} is and is not in η_{i+1}. If it is, then we can check if v_iv_{i-1} is a blocking edge with respect to η_{i+1} in P_{i+1}. Otherwise, v_iv_{i+1} is a blocking edge with respect to η_{i+1} in P_{i+1}. All these possibilities can be taken care by appropriately defining the table entries. In the table, however, we do not need to store the whole matching. We only need to remember the matching partner of vertices, such as v_{i-1}, as that will help in checking if v_iv_{i+1} is a blocking edge. We can similarly argue for a cycle in G.

Next, we would like to point out that our hardness results, that is, Theorems 1, 2, and 3 hold even when the preference lists of the vertices in each side of the partition respect a master ordering of vertices i.e., the relative ordering of the vertices in a preference list is same as that of a fixed ordering of all the vertices on the other side. We discuss it in details in the full version of the paper.

Future work. We conclude the paper with a few directions for further research.

- In certain scenarios, the "satisfaction" of the agents (there exist several measures such as *egalitarian*, *sex-equal*, *balance*) might be of importance. Then, it might be of interest to study the tradeoff between t and k, tradeoff between egalitarian/sex-equal/balance cost and k.
- The formulation of ASM can be generalized to the case where the input contains a utility function on the edges and the objective is to maximize the value of a solution matching subject to this function.

References

1. D. J. Abraham, P. Biró, and D. F. Manlove. "Almost stable" matchings in the roommates problem. In *International Workshop on Approximation and Online Algorithms (WG)*, pages 1–14, 2005.
2. P. Biró, T. Fleiner, R. W. Irving, and D. F. Manlove. The college admissions problem with lower and common quotas. *Theoretical Computer Science*, 411(34-36):3136–3153, 2010.
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. P. Biró, D. F. Manlove, and E. J. McDermid. "Almost stable" matchings in the roommates problem with bounded preference lists. *Theoretical Computer Science*, 432:10–20, 2012.
5. J. Chen, D. Hermelin, M. Sorge, and H. Yedidsion. How hard is it to satisfy (almost) all roommates? In *International Colloquium on Automata, Languages, and Programming, (ICALP)*, pages 35:1–35:15, 2018.
6. J. Chen, P. Skowron, and M. Sorge. Matchings under preferences: Strength of stability and trade-offs. In *ACM Conference on Economics and Computation (EC)*, pages 41–59, 2019.
7. Á. Cseh and D. F. Manlove. Stable marriage and roommates problems with restricted edges: Complexity and approximability. *Discrete Optimization*, 20:62–89, 2016.
8. M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.
9. R. Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
10. R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.
11. M. R. Fellows, F. V. Fomin, D. Lokshtanov, F. A. Rosamond, S. Saurabh, and Y. Villange. Local search: Is brute-force avoidable? *Journal of Computer and System Sciences*, 78(3):707–719, 2012.
12. Michael R. Fellows, Danny Hermelin, Frances A. Rosamond, and Stéphane Vialette. On the parameterized complexity of multiple-interval graph problems. *Theor. Comput. Sci.*, 410(1):53–61, 2009.
13. F. V. Fomin, D. Lokshtanov, F. Panolan, and S. Saurabh. Efficient computation of representative families with applications in parameterized and exact algorithms. *Journal of the ACM*, 63(4):29:1–29:60, 2016.
14. D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

15 D. Gusfield and R. W. Irving. *The stable marriage problem: structure and algorithms*. MIT press, 1989.

16 R. W. Irving and D. Manlove. Finding large stable matchings. *ACM Journal of Experimental Algorithmics*, 2009.

17 H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack problems*. Springer, 2004.

18 S. Khuller, R. Bhatia, and R. Pless. On local search and placement of meters in networks. *SIAM journal on computing*, 32(2):470–487, 2003.

19 D. E. Knuth. Marriages stables. *Technical report*, 1976.

20 A. Krokhin and D. Marx. On the hardness of losing weight. *ACM Transactions on Algorithms*, 8(2):1–18, 2012.

21 D. Manlove. *Algorithmics of matching under preferences*, volume 2. World Scientific, 2013.

22 D. F. Manlove, I. McBride, and J. Trimble. "Almost-stable" matchings in the Hospitals/Residents problem with couples. *Constraints*, 22(1):50–72, 2017.

23 D. Marx. Local search. *Parameterized Complexity News*, 3:7–8, 2008.

24 D. Marx. Searching the k-change neighborhood for TSP is W[1]-hard. *Operations Research Letters*, 36(1):31–36, 2008.

25 D. Marx and I. Schlotter. Parameterized complexity and local search approaches for the stable marriage problem with ties. *Algorithmica*, 58(1):170–187, 2010.

26 D. Marx and I. Schlotter. Stable assignment with couples: Parameterized complexity and local search. *Discrete Optimization*, 8(1):25–40, 2011.

27 S. Micali and V. V. Vazirani. An $O(\sqrt{|V|}|E|)$ algorithm for finding maximum matching in general graphs. In *Foundations of Computer Science (FOCS)*, pages 17–27, 1980.

28 R. Neidermeier. *Invitation to fixed-parameter algorithms*. Springer, 2006.

29 Krzysztof Pietrzak. On the parameterized complexity of the fixed alphabet shortest common supersequence and longest common subsequence problems. *J. Comput. Syst. Sci.*, 67(4):757–771, 2003.

30 A. E. Roth. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, 92(6), 1984.

31 A. E. Roth. On the allocation of residents to rural hospitals: A general property of two-sided matching markets. *Econometrica: Journal of the Econometric Society*, 54(2):425–427, 1986.

32 S. Szeider. The parameterized complexity of k-flip local search for sat and max sat. *Discrete Optimization*, 8(1):139–145, 2011.

33 El-Ghazali Talbi. *Metaheuristics: From design to implementation*, volume 74. John Wiley & Sons, 2009.

34 K. Tomoeda. Finding a stable matching under type-specific minimum quotas. *Journal of Economic Theory*, 176:81–117, 2018.