Theoretical Computer Science 883 (2021) 19–43

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Balanced stable marriage: How close is close enough?

Sushmita Gupta [a,*,1], Sanjukta Roy [a,*,**], Saket Saurabh [b,a,*,2], Meirav Zehavi [c,*,3]

[a] The Institute of Mathematical Sciences, HBNI, Chennai, India
[b] University of Bergen, Bergen, Norway
[c] Ben-Gurion University, Beersheba, Israel

A R T I C L E   I N F O

Article history:
Received 22 January 2019
Received in revised form 31 January 2021
Accepted 8 May 2021
Available online 17 June 2021
Communicated by P. Krysta

Keywords:
Keywords balanced stable marriage
Parameterized complexity
Kernelization

A B S T R A C T

Balanced Stable Marriage (BSM) is a central optimization version of the classic Stable Marriage (SM) problem. We study BSM from the viewpoint of Parameterized Complexity. Informally, the input of BSM consists of n men, n women, and an integer k. Each person a has a (sub)set of acceptable partners, $\mathcal{A}(a)$, whom a ranks strictly; we use $p_a(b)$ to denote the position of $b \in \mathcal{A}(a)$ in a's preference list. The objective is to decide whether there exists a stable matching μ such that $\mathsf{balance}(\mu) \triangleq \max\{\sum_{(m,w)\in\mu} p_m(w), \sum_{(m,w)\in\mu} p_w(m)\} \leq k$. In SM, all stable matchings match the same set of agents, $A^\star$ which can be computed in polynomial time. As $\mathsf{balance}(\mu) \geq \frac{|A^\star|}{2}$ for any stable matching μ, BSM is trivially fixed-parameter tractable (FPT) with respect to k. Thus, a natural question is whether BSM is FPT with respect to $k - \frac{|A^\star|}{2}$. With this viewpoint in mind, we draw a line between tractability and intractability in relation to the target value. This line separates additional natural parameterizations higher/lower than ours (e.g., we automatically resolve the parameterization $k - \frac{|A^\star|}{2}$). The two extreme stable matchings are the man-optimal μ_M and the woman-optimal μ_W. Let $O_M = \sum_{(m,w)\in\mu_M} p_m(w)$, and $O_W = \sum_{(m,w)\in\mu_W} p_w(m)$. In this work, we prove that

- BSM parameterized by $t = k - \min\{O_M, O_W\}$ admits (1) a kernel where the number of people is linear in t, and (2) a parameterized algorithm whose running time is single exponential in t.
- BSM parameterized by $t = k - \max\{O_M, O_W\}$ is W[1]-hard.

* Corresponding authors.
** Contact author.
E-mail addresses: sushmitagupta@imsc.res.in (S. Gupta), sanjuktaroy.mail@gmail.com (S. Roy), saket@imsc.res.in (S. Saurabh), meiravze@bgu.ac.il (M. Zehavi).

[1] Received funding from SERB-Starting Research Grant (SRG/2019/001870).
[2] 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.
[3] Received funding from ISF grant (no. 1176/18) and BSF grant (no. 2018302).

https://doi.org/10.1016/j.tcs.2021.05.015
0304-3975/© 2021 Published by Elsevier B.V.

1. Introduction

Over the last two decades, Parameterized Complexity has evolved to be a central field of research in theoretical computer science. However, the scope of this field had a strong focus on applications to NP-hard optimization problems *on graphs*. There is no inherent reason why this should be the case. Indeed, the main idea of Parameterized Complexity is very general—measure running time in terms of both input size and a parameter that captures structural properties of the input instance. The idea of a multivariate analysis of algorithms holds the potential to address the need for a framework for refined algorithm analysis for all kinds of problems across all domains and subfields of computer science. Recently, techniques in Parameterized Complexity were successfully applied in the area of Computational Social Choice Theory. In particular, Voting has become a subject of intensive study from the viewpoint of Parameterized Complexity (for a few examples, see [1, Chapters 10,11] and [3]; for more information on the current state-of-the-art, we refer to excellent surveys such as [4,7,8,14]. However, Voting is only one topic under the rich umbrella of Computational Social Choice. Parameterized analysis of other topics has been few and far between. In the past few months, a collective effort to study matching under preferences through the lens of Parameterized Complexity was initiated [6,9,19,20,24,25,32,33,36,37,40,41]. Along with [6,9], our work has initiated the study of *kernelization* in the context of matchings under preferences. To the best of our knowledge, our work is the first to introduce "above-guarantee parameterization" to this topic.

In Matching under Preferences, a matching is an allocation (or assignment) of agents to resources that satisfies some predefined criterion of compatibility/acceptability. Here, (arguably) the best known model is the *two-sided model*, where the agents on one side are referred to as *men*, and the agents on the other side are referred to as *women*. A few illustrative examples of real life situations where this model is relevant include matching residents to hospitals, students to colleges, kidney patients to donors and users to servers in a distributed Internet service. At the heart of all of these applications lies the fundamental Stable Marriage problem. In particular, the Nobel Prize in Economics was awarded to Shapley and Roth in 2012 "for the theory of stable allocations and the practice of market design." Moreover, several books have been dedicated to the study of Stable Marriage as well as optimization variants of this classical problem such as the Egalitarian Stable Marriage, Minimum Regret Stable Marriage, Sex-Equal Stable Marriage and Balanced Stable Marriage problems [21,22,31,35].

The input of Stable Marriage consists of a set of men, M, and a set of women, W. Each person a has a set of *acceptable partners*, $\mathcal{A}(a)$, who are subjectively ranked by a in a strict order. Consequently, each person a has a so-called *preference list*, where $p_a(b)$ is the position of $b \in \mathcal{A}(a)$ in a's preference list. Without loss of generality, it is assumed that if a person a ranks a person b, then the person b ranks the person a as well. The sets of preference lists of the men and the women are denoted by $\mathcal{L}_M$ and $\mathcal{L}_W$, respectively. In this context, we say that a pair of a man and a woman, (m, w), is an *acceptable pair* if both $m \in \mathcal{A}(w)$ and $w \in \mathcal{A}(m)$. Accordingly, the notion of a *matching* refers to a matching between men and women, where two people matched to one another form an acceptable pair. Roughly speaking, the goal of the Stable Marriage problem is to *find* a matching that is *stable* in the following sense: there should not exist two people who prefer being matched to one another over their current "status". More precisely, a matching μ is said to be stable if it does not have a *blocking pair*, which is an acceptable pair (m, w) such that **(i)** either m is unmatched by μ or $p_m(w) < p_m(\mu(m))$, and **(ii)** either w is unmatched by μ or $p_w(m) < p_w(\mu(w))$. Here, the notation $\mu(a)$ indicates the person to whom μ matches the person a. Note that a person always prefers being matched to an acceptable partner over being unmatched.

The seminal paper [16] by Gale and Shapely on stable matchings shows that given an instance of Stable Marriage, a stable matching necessarily *exists*, but it is not necessarily unique. In fact, for a given instance of Stable Marriage, there can be an *exponential* number of stable matchings, and they should be viewed as a *spectrum* where the two extremes are known as the *man-optimal stable matching* and the *woman-optimal stable matching*. Formally, the man-optimal stable matching, denoted by μ_M, is a stable matching such that every stable matching μ satisfies the following condition: every man m is either unmatched by both μ_M and μ or $p_m(\mu_M(m)) \leq p_m(\mu(m))$. The woman-optimal stable matching, denoted by μ_W, is defined analogously. These two extremes, which give the best possible solution for one party at the expense of the other party, always exist and can be computed in polynomial time [16]. Naturally, it is desirable to analyze matchings that lie somewhere in the middle. Here, the quantity $p_a(\mu(a))$ is the "dissatisfaction" of a in a matching μ, where a smaller value signifies a smaller amount of dissatisfaction. The most well-known measures are as follows:

- μ is *globally desirable* if it minimizes $\sum_{(m,w)\in\mu}(p_m(w) + p_w(m))$, called the *egalitarian stable matching*;
- μ is *minimum regret* if it minimizes $\max_{(m,w)\in\mu}\{\max\{p_m(w), p_w(m)\}\}$, called the *minimum regret stable matching*;
- μ is *fair towards both sides* if it minimizes $|\sum_{(m,w)\in\mu} p_m(w) - \sum_{(m,w)\in\mu} p_w(m)|$, called the *sex-equal stable matching*;
- μ is *desirable by both sides* if it minimizes $\max\{\sum_{(m,w)\in\mu} p_m(w), \sum_{(m,w)\in\mu} p_w(m)\}$, called the *balanced stable matching*.

Each notion above leads to a natural, *different* well-studied optimization problem (see Related Work). We focus on the NP-hard Balanced Stable Marriage (BSM) problem, where the objective is to find a stable matching μ that minimizes

$$\mathsf{balance}(\mu) = \max\{\sum_{(m,w)\in\mu} p_m(w), \sum_{(m,w)\in\mu} p_w(m)\}.$$

This problem was introduced in the influential work of Feder [15] on stable matchings, and was shown to be NP-hard and admitting a 2-approximation algorithm.

Our Contribution. Above-guarantee parameterization is a topic of extensive study in Parameterized Complexity [10]. We introduce two "above-guarantee parameterizations" of BSM. Consider the minimum value O_M (O_W) of the total dissatisfaction of men (women) realizable by a stable matching. Formally, $O_M = \sum_{(m,w)\in\mu_M} p_m(w)$, and $O_W = \sum_{(m,w)\in\mu_W} p_w(m)$, where μ_M (μ_W) is the man-optimal (woman-optimal) stable matching. Denote $\mathsf{Bal} = \min_{\mu\in\mathsf{SM}} \mathsf{balance}(\mu)$, where SM is the set of all stable matchings. An input integer k would indicate that our objective is to decide whether $\mathsf{Bal} \le k$. In our first parameterization, the parameter is $k - \min\{O_M, O_W\}$, and in the second one, it is $k - \max\{O_M, O_W\}$.

Above-Min Balanced Stable Marriage (Above-Min BSM)
Input: An instance $(M, W, \mathcal{L}_M, \mathcal{L}_W)$ of Balanced Stable Marriage, and a non-negative integer k.
Question: Is $\mathsf{Bal} \le k$?
Parameter: $t = k - \min\{O_M, O_W\}$.

Above-Max Balanced Stable Marriage (Above-Max BSM)
Input: An instance $(M, W, \mathcal{L}_M, \mathcal{L}_W)$ of Balanced Stable Marriage, and a non-negative integer k.
Question: Is $\mathsf{Bal} \le k$?
Parameter: $t = k - \max\{O_M, O_W\}$.

Choice of parameters: Note that the least dissatisfaction the party of men can hope for (call it *minimum dissatisfaction*) is O_M, and the least dissatisfaction the party of women can hope for (also call it minimum dissatisfaction) is O_W. First, consider the parameter $t = k - \min\{O_M, O_W\}$. Whenever we have a solution such that the amounts of dissatisfaction of *both* parties are *close enough* to the minimum, this parameter is small. (When the parameter is small, we cannot simply pick μ_M or μ_W since $\mathsf{balance}(\mu_M)$ and $\mathsf{balance}(\mu_W)$ can be *arbitrarily larger* than $\min\{O_M, O_W\}$.) Indeed, the closer the dissatisfaction of both parties to the minimum (which is exactly the case where both parties would find the solution desirable), the smaller the parameter is. The smaller the parameter is, the faster is a parameterized algorithm. In this above-guarantee parameterization, the guarantee value (i.e., $\min\{O_M, O_W\}$) is already quite high—for example, our parameter is significantly smaller than $k - n'$, where n' is the number of men (or women) matched by a stable matching, since $k - \min\{O_M, O_W\}$ is (i) never larger than $k - n'$, and (ii) can be *arbitrarily smaller* than $k - n'$, e.g. $k - n'$ can be of the magnitude of $\mathcal{O}(n)$ while our parameter is 0.

Since we are taking the *minimum* of $\{O_M, O_W\}$, we need the dissatisfaction of *both* parties to be close to optimum in order to have a small parameter. As we are able to show that BSM is FPT with respect to this parameter, it is very natural to next examine the case where we take the *maximum* of $\{O_M, O_W\}$. In this case, the closer the dissatisfaction of *at least one* party to the optimum, the smaller the parameter is. In other words, now to have a small parameter, the demand from a solution is weaker than before. In the vocabulary of Parameterized Complexity, it is said that the parameterization by $t = k - \max\{O_M, O_W\}$ is "above a higher guarantee" than the parameterization by $t = k - \min\{O_M, O_W\}$, since it is *always* the case that $\max\{O_M, O_W\} \ge \min\{O_M, O_W\}$. Interestingly, as we show in this paper, the parameterization by $k - \max\{O_M, O_W\}$ results in a problem that is W[1]-hard. Hence, the complexities of the two parameterizations behave very differently. We remark that in Parameterized Complexity, it is *not at all* the rule that when one takes an "above a higher guarantee" parameterization, the problem would become W[1]-hard, as can be evidenced by the Vertex Cover problem, the classical above-guarantee parameterizations in this field, for which three distinct above-guarantee parameterizations yielded FPT algorithms [11,18,34,44]. Overall, our results draw a nontrivial line between tractability and intractability of above-guarantee parameterization of BSM.

Finally, the three main theorems that we establish in this study are as follows.

Theorem 1. Above-Max BSM *is* W[1]-hard.

Theorem 2. Above-Min BSM *admits a kernel that has at most* $3t$ *men, at most* $3t$ *women, and such that each person has at most* $2t + 1$ *acceptable partners.*

Theorem 3. Above-Min BSM *can be solved in time*[4] $\mathcal{O}^*(8^t)$.

Our techniques: The overview of the reduction we develop to prove Theorem 1 is presented followed by the formal analysis in Section 3. The proof of Theorem 2 is based on the introduction and analysis of a variant of Above-Min BSM which we will call Above-Min FBSM Our kernelization algorithm consists of several phases, each simplifying a different aspect of Above-Min BSM, and shedding light on structural properties of the Yes-instances of this problem. We stress that the design and

4 $\mathcal{O}^*$ hides the factors polynomial in input size.

order of our reduction rules are very carefully tailored to ensure correctness. For example, it is tempting to execute Step 2 before Step 1 (see the outline in Section 4.1), but this is simply incorrect as it alters the set of stable matchings. The reduction rules are easy to express for this functional variant of Above-Min BSM. Hence, we choose to define the reduction rules for this functional variant of Above-Min BSM instead of Above-Min BSM directly. Note that Theorem 2 already implies that Above-Min BSM is FPT. To obtain a parameterized algorithm whose running time is single exponential in the parameter (Theorem 3), we utilize the method of bounded search trees on top of our kernel in Section 5.

Related work. The problem of finding an egalitarian stable matching (defined on page 20), called Egalitarian Stable Marriage (ESM), is known to be solvable in polynomial time due to Irving et al. [28]. If preference lists have so called *ties* or agents do not have genders (i.e., in *roommates* setting), then ESM is NP-hard [27]. ESM is one of the most well studied topics in Matching under Preferences (see [22,31,35]), and the survey of these results is outside the scope of this paper. ESM does not distinguish between men and women, and therefore it does not fit scenarios where it is necessary to differentiate between the individual satisfaction of each party. In such scenarios, Sex-Equal Stable Marriage (SESM) and BSM come into play.

In the SESM problem, the objective is to find a sex-equal stable matching, a stable matching that is fair towards both sides by minimizing the difference between their individual amounts of satisfaction. Unlike EGM, the SESM problem is known to be NP-hard [30]. At first sight, the balance measure might seem conceptually similar to the sex-equal measure, but in fact, the two measures are quite different. Indeed, BSM does not attempt to find a stable matching that is fair, but one that is desirable by both sides. In other words, BSM examines the amount of dissatisfaction of each party *individually*, and attempts to minimize the worse one among the two. BSM models the scenario where each party is selfish in the sense that it desires a matching where its own dissatisfaction is minimized, irrespective of the dissatisfaction of the other party, and our goal is to find a matching desirable by both parties by ensuring that each individual amount of dissatisfaction does not exceed some threshold. However, even if parties are not selfish, a balanced stable matching may be preferable over a sex-equal one—indeed, in some situations, a sex-equal stable matching may fit the old phrase "the sorrow of many is the comfort of fools", as it can achieve equality not by making one party better off at the expense of the other, but simply by making both parties worse off. Thus, although the amount of literature devoted to SESM is much greater than the one about BSM, we find BSM at least as important as SESM. The survey of results about SESM is outside the scope of this paper, but let us remark that in some situations, the minimization of $\mathsf{balance}(\mu)$ may indirectly also minimize $|\delta(\mu)|$ where $\delta(\mu) = \sum_{(m,w)\in\mu} p_m(\mu(m)) - \sum_{(m,w)\in\mu} p_w(\mu(w))$, but in other situations, this may not be the case. Indeed, McDermid [38] constructed a family of instances where there does *not* exist any matching that is both a sex-equal stable matching and a balanced stable matching (the construction is also available in the book [35]).

Manlove discusses that BSM also admits a $(2-1/\ell)$-approximation algorithm where ℓ is the maximum size of a set of acceptable partners [35]. O'Malley [42] phrased the BSM problem in terms of constraint programming. The parameterized complexity of BSM with respect to treewidth parameterizations (of graphs associated with the problem) was studied in [20]. Recently, McDermid and Irving [39] expressed explicit interest in the design of fast exact exponential-time algorithms for BSM. McDermid and Irving [39] showed that SESM is NP-hard even if it is only necessary to decide whether the target $\Delta = \min_{\mu\in\mathsf{SM}} |\delta(\mu)|$ is 0 or not [39], where SM is the set of all stable matchings of the instance. In particular, this means that SESM is not only W[1]-hard with respect to Δ, but it is even paraNP-hard with respect to this parameter.[5] In the case of BSM, however, the problem is trivially FPT with respect to the target $\mathsf{Bal} = \min_{\mu\in\mathsf{SM}} \mathsf{balance}(\mu)$.

In the framework of Parameterized Complexity, Chen et al. [9] showed that deciding whether there exists a matching with at most k blocking pairs in the roommates setting is W[1]-hard. In addition to BSM, Gupta et al. [20] also considered treewidth parameterizations of other optimization variants of Stable Marriage. Mnich and Schlotter [41] inspected the parameterized complexity of an optimization variant of Stable Marriage where blocking pairs were allowed to exist in order to ensure that some agents will be matched by the output matching. Meeks and Rastegari [40] introduced a parameterization regarding "agent types" and studied several optimization variants of Stable Marriage with respect to it. Marx and Schlotter [36] studied size optimization variants of Stable Marriage in the presence of ties, where the parameters are the total length, maximum length, and number of ties. In addition, Marx and Schlotter [37] studied a generalization of Stable Marriage, called Hospitals/Residents with Couples, with the number of couples as a parameter. Other studies of parameterized complexity of problems (less closely) related to Stable Marriage are of the Group Activity Selection problem [19,24,25] and of the Stable Invitations problem [32]. Some works that are not directly related to our work but study the parameterized complexity of problems that involve restricted preference lists are [2,5,23,26,29].

2. Preliminaries

Throughout the paper, whenever the instance $\mathcal{I}$ of BSM under discussion is not clear from context or we would like to put emphasis on $\mathcal{I}$, we add "$(\mathcal{I})$" to the appropriate notation. For example, we use the notation $t(\mathcal{I})$ rather than t. When we would like to refer to the balance of a stable matching μ in a specific instance $\mathcal{I}$, we would use the notation $\mathsf{balance}_{\mathcal{I}}(\mu)$. A matching is called a *perfect matching* if it matches every person (to some other person).

[5] If a parameterized problem is NP-hard even when the value of the parameter is a fixed constant (that is, independent of the input), then the problem is said to be paraNP-hard.

A Functional variant of Stable Marriage. To obtain our kernelization algorithm for Above-Min BSM, it will be convenient to work with, as explained below, a "functional" definition of preferences, resulting in a "functional" variant of this problem which we call Above-Min FBSM. Here, instead of the sets of preferences lists $\mathcal{L}_M$ and $\mathcal{L}_W$, the input consists of sets of preference *functions* $\mathcal{F}_M$ and $\mathcal{F}_W$, where $\mathcal{F}_M$ replaces $\mathcal{L}_M$ and $\mathcal{F}_W$ replaces $\mathcal{L}_W$. Specifically, every person $a \in M \cup W$ has an injective (one-to-one) function $f_a : \mathcal{A}(a) \to \mathbb{N}$, called a *preference function*. Similar models are known in literature [43]. Intuitively, a lower function value corresponds to a higher preference. The preference functions are injective because each person's preference list is a strict order of his/her acceptable partners. Consequently, every preference function defines a total ordering over a set of acceptable partners. Note that all of the definitions presented in the introduction extend to our functional variant in the natural way. For the sake of formality, we will specify the necessary adaptations. Our kernelization algorithm for Above-Min BSM works as follows: Firstly, we create an instance of Above-Min FBSM, secondly, we run the kernelization algorithm for Above-Min FBSM to obtain a kernel for our instance; and finally, we transform that kernel into a kernel for the original instance of Above-Min BSM.

We adapt standard definitions presented for Above-Min BSM to the case where preferences are specified by functions rather than lists. The input of the Functional Stable Marriage problem consists of a set of men, M, and a set of women, W. Each person a has a set of *acceptable partners*, denoted by $\mathcal{A}(a)$, and an injective function $f_a : \mathcal{A}(a) \to \mathbb{N}$ called a *preference function*. Without loss of generality, it is assumed that if a person a belongs to the set of acceptable partners of a person b, then the person b belongs to the set of acceptable partners of the person a. The set of preference functions of the men and the women are denoted by $\mathcal{F}_M$ and $\mathcal{F}_W$, respectively. A pair of a man and a woman, (m, w), is an *acceptable pair* if both $m \in \mathcal{A}(w)$ and $w \in \mathcal{A}(m)$. Accordingly, the notion of a *matching* refers to a set of mutually disjoint pairs comprising of a man and a woman, where each (man, women) pair forms an acceptable pair. A matching μ is *stable* if it does not have a *blocking pair*, which is an acceptable pair (m, w) such that **(i)** either m is unmatched by μ or $f_m(w) < f_m(\mu(m))$, and **(ii)** either w is unmatched by μ or $f_w(m) < f_w(\mu(w))$. The goal of the Functional Stable Marriage problem is to find a stable matching.

The man-optimal stable matching, denoted by μ_M, is a stable matching such that every stable matching μ satisfies the following condition: every man m is either unmatched by both μ_M and μ or $f_m(\mu_M(m)) \leq f_m(\mu(m))$. The woman-optimal stable matching, denoted by μ_W, is defined analogously. Given a stable matching μ, define $\mathsf{balance}(\mu) = \max\{\sum_{(m,w)\in\mu} f_m(w), \sum_{(m,w)\in\mu} f_w(m)\}$. Moreover, $\mathsf{Bal} = \min_{\mu\in\mathsf{SM}} \mathsf{balance}(\mu)$, where SM is the set of all stable matchings, $O_M = \sum_{(m,w)\in\mu_M} f_m(w)$, and $O_W = \sum_{(m,w)\in\mu_W} f_w(m)$. Finally, Above-Min FBSM is defined as follows.

> Above-Min Functional Balanced Stable Marriage (Above-Min FBSM)
> **Input:** An instance $(M, W, \mathcal{F}_M, \mathcal{F}_W)$ of Functional Balanced Stable Marriage, and a non-negative integer k.
> **Question:** Is $\mathsf{Bal} \leq k$?
> **Parameter:** $t = k - \min\{O_M, O_W\}$.

Clearly, it is straightforward to turn an instance of Above-Min BSM into an equivalent instance of Above-Min FBSM as stated in the following observation.

Observation 2.1. *Let $\mathcal{I} = (M, W, \mathcal{L}_M, \mathcal{L}_W, k)$ be an instance of* Above-Min BSM*. For each $a \in M \cup W$, define $f_a : \mathcal{A}(a) \to \mathbb{N}$ by setting $f_a(b) = p_a(b)$ for all $b \in \mathcal{A}(a)$. Then, $\mathcal{I}$ is a* Yes*-instance of* Above-Min BSM *if and only if $(M, W, \mathcal{F}_M = \{f_m\}_{m\in M}, \mathcal{F}_W = \{f_w\}_{w\in W}, k)$ is a* Yes*-instance of* Above-Min FBSM.

2.1. Known results

Here, we state several classical results, which were originally presented in the context of Stable Marriage. By their original proofs, these results also hold in the context of Functional Stable Marriage. To be more precise, given an instance of Functional Stable Marriage, we can construct an equivalent instance of Stable Marriage, by ranking the acceptable partners in the order of their function values, where a smaller value implies a higher preference. The instances are equivalent in the sense that they give rise to the exact same set of stable matchings. Hence, all the structural results about stable matchings in the usual setting (modeled by strict preference lists) apply to the generalized setting, modeled by injective functions.

Proposition 2.1 *([16]). For any instance of* Stable Marriage *(or* Functional Stable Marriage*), there exist a man-optimal stable matching, μ_M, and a woman-optimal stable matching, μ_W, and both μ_M and μ_W can be computed in time $\mathcal{O}(m)$ where m is the number of acceptable pairs.*

The following powerful proposition is known as the Rural-Hospital Theorem (and notably does not hold for instances that have ties).

Proposition 2.2 *([17] Theorem 1). Given an instance of* Stable Marriage *(or* Above-Min FBSM*), the set of men and women that are matched is the same for all stable matchings.*

We further need a proposition regarding the man-optimal and woman-optimal stable matchings that implies Proposition 2.2 [17].

Proposition 2.3 *([22]). For any instance of* Stable Marriage *(or* Functional Stable Marriage*), every stable matching μ satisfies the following conditions: every woman w is either unmatched by both μ_M and μ or $p_w(\mu_M(w)) \geq p_w(\mu(w))$, and every man m is either unmatched by both μ_W and μ or $p_m(\mu_W(m)) \geq p_m(\mu(m))$.*

2.2. Parameterized complexity

A *parameterization* of an optimization problem Π is a pair (I,k) where I is an instance of Π and k is a non-negative integer. For example, an instance of Clique parameterized by the solution size is a pair (G,k), where G is an undirected graph encoded as a string over a fixed finite alphabet Σ, and k is a positive integer. The pair (G,k) is a Yes-instance of the parameterized problem if and only if the string G correctly encodes an undirected graph and that graph has a clique on k vertices.

For our purposes, we need to recall three central notions that define the parameterized complexity of a parameterized problem. The first one is the notion of a *kernel*. Here, a parameterized problem is said to admit a *kernel* of size $f(k)$ for some function f that depends *only* on k if there exists a polynomial-time algorithm, called a *kernelization algorithm*, that translates any input instance into an equivalent instance of the same problem whose size is bounded by $f(k)$ and such that the value of the parameter does not increase. In case the function f is polynomial in k, the problem is said to admit a *polynomial kernel*. Hence, kernelization is a mathematical concept that aims to analyze the power of preprocessing procedures in a formal, rigorous manner. The second notion that we use is the one of *fixed-parameter tractability (*FPT*)*. Here, a parameterized problem Π is said to be FPT if there is an algorithm that solves it in time $f(k) \cdot |I|^{\mathcal{O}(1)}$, where $|I|$ is the size of the input and f is a function that depends only on k. Such an algorithm is called a *parameterized algorithm*. In other words, the notion of FPT signifies that it is not necessary for the combinatorial explosion in the running time of an algorithm for Π to depend on the input size, but it can be confined to the parameter k. Finally, we recall that Parameterized Complexity also provides tools to refute the existence of polynomial kernels and parameterized algorithms for certain problems (under plausible complexity-theoretic assumptions), in which context the notion of W[1]-hardness is a central one. It is widely believed that a problem that is W[1]-hard is unlikely to be FPT, and we refer the reader to the books [10,13] for more information on this notion in particular, and on Parameterized Complexity in general.

While designing our kernelization algorithm, we might be able to determine whether the input instance is a Yes-instance or a No-instance. For the sake of clarity, in the first case, we simply return Yes, and in the second case, we simply return No. To properly comply with the definition of a kernel, the return of Yes and No should be interpreted as the return of a trivial Yes-instance and a trivial No-instance, respectively. Here, a trivial Yes-instance can be the one in which $M = W = \emptyset$ and $k = 0$, where the only stable matching is the one that is empty and whose balance is 0, and a trivial No-instance can be the one where $M = \{m\}$, $W = \{w\}$, $\mathcal{A}(m) = \{w\}$, $\mathcal{A}(w) = \{m\}$ and $k = 0$.

Reduction Rule. To design our kernelization algorithm, we rely on the notion of a *reduction rule*. A reduction rule is a polynomial-time procedure that replaces an instance $(\mathcal{I}, k)$ of a parameterized problem Π by a new instance $(\mathcal{I}', k')$ of Π. Roughly speaking, a reduction rule is useful when the instance $\mathcal{I}'$ is in some sense "simpler" than the instance $\mathcal{I}$. In particular, it is desirable to ensure that $k' \leq k$. The rule is said to be *safe* if $(\mathcal{I}, k)$ is a Yes-instance if and only if $(\mathcal{I}', k')$ is a Yes-instance. The reduction rules designed in this paper will be applied repeatedly, as long as they are applicable, in the order they are described.

3. Hardness

In this section, we prove Theorem 1. For this purpose, we give a reduction from a W[1]-hard problem, Clique [12]. Thus, to prove Theorem 1, it is sufficient to prove the next lemma. First, we define Clique formally.

> Clique
> **Input:** A graph $G = (V, E)$, and a positive integer k.
> **Question:** Does G contain a complete graph on k vertices?
> **Parameter:** k.

Lemma 3.1. *Given an instance $\mathcal{I} = (G = (V, E), k)$ of* Clique*, an equivalent instance $\widehat{\mathcal{I}} = (M, W, \mathcal{L}_M, \mathcal{L}_W, \widehat{k})$ of* Above-Max BSM *with parameter $t = 6k + 3k(k-1)$ can be constructed in time $f(k) \cdot |\mathcal{I}|^{\mathcal{O}(1)}$ for some function f.*

The goal is to construct (in "FPT time") an instance $\widehat{\mathcal{I}} = (M, W, \mathcal{L}_M, \mathcal{L}_W, \widehat{k})$ of Above-Max BSM. The following subsections contain an informal explanation of the intuition underlying the gadget construction followed by the formal description and the analysis.

Table 1
Preference list for every man $m \in M$, and thus constructing $\mathcal{L}_M$. Here, $[X]$ denotes a fixed permutation of X.

M	Preference list	
m_v^1	$w_v^1, \widetilde{w}^1, \widetilde{w}^2, w_v^2$	for all $m_v^1 \in M_V$
m_v^2	$w_v^2, \widetilde{w}^1, \widetilde{w}^2, w_v^1$	for all $m_v^2 \in M_V$
$m_{\{u,v\}}^1$	$w_{\{u,v\}}^1, w_u^1, w_v^1, w_{\{u,v\}}^2$	for all $m_{\{u,v\}}^1 \in M_E$ where* $u < v$
$m_{\{u,v\}}^2$	$w_{\{u,v\}}^2, w_u^2, w_v^2, w_{\{u,v\}}^1$	for all $m_{\{u,v\}}^2 \in M_E$ where $u < v$
$\widetilde{m}^i$	$\widetilde{w}^i, w_{e_1}^1, w_{e_2}^1, \ldots, w_{e_{\|E\|}}^1, w_{e_1}^2, w_{e_2}^2, \ldots, w_{e_{\|E\|}}^2, [X]$	for all $\widetilde{m}^i \in \widetilde{M}$ such that $i \leq \|V\|\|E\|$, $X = \{w_v^j \in W_V : \widetilde{m}^i \in \mathcal{A}(w_v^j), j \in \{1,2\}\}$
$\widetilde{m}^i$	$\widetilde{w}^i$	for all $\widetilde{m}^i \in \widetilde{M}$ such that $i > \|V\|\|E\|$
m^*	$\widetilde{w}^1, \widetilde{w}^2, \ldots \widetilde{w}^\delta, w^*$	

* Recall that we have defined an order on V.

Reduction. Let $\mathcal{I} = (G = (V, E), k)$ be some instance of Clique. We select arbitrary orders on V and E, and accordingly we denote $V = \{v_1, v_2, \ldots, v_{|V|}\}$ and $E = \{e_1, e_2, \ldots, e_{|E|}\}$.

First, to construct the sets M and W, we define three pairwise-disjoint subsets of M, called M_V, M_E and $\widetilde{M}$, and three pairwise-disjoint subsets of W, called W_V, W_E and $\widetilde{W}$. Then, we set $M = M_V \uplus M_E \uplus \widetilde{M} \uplus \{m^*\}$ and $W = W_V \uplus W_E \uplus \widetilde{W} \uplus \{w^*\}$, where m^* and w^* denote a new man and a new woman, respectively.

- $M_V = \{m_v^i : v \in V, i \in \{1,2\}\}$; $W_V = \{w_v^i : v \in V, i \in \{1,2\}\}$.
- $M_E = \{m_e^i : e \in E, i \in \{1,2\}\}$; $W_E = \{w_e^i : e \in E, i \in \{1,2\}\}$.
- Let $\delta = 2(|V| + |E| + |V||E| + |V||E|^2) - k(4 + 4k + 2|E| + (k-1)|V||E|)$.
  Then, $\widetilde{M} = \{\widetilde{m}^i : i \in \{1,2,\ldots,\delta\}\}$ and $\widetilde{W} = \{\widetilde{w}^i : i \in \{1,2,\ldots,\delta\}\}$.

Note that $|M| = |W|$. We remark that in what follows, we assume w.l.o.g. that $\delta \geq 0$ and $|V| > k + k(k-1)/2$, else the size of the input instance $\mathcal{I}$ of Clique is bounded by a function of k and can therefore, by using brute-force, be solved in FPT time.

Roughly speaking, each pair of men, m_v^1 and m_v^2, represents a vertex, and we aim to ensure that either both men will be matched to their best partners (in the man-optimal stable matching) or both men will be matched to other partners (where there would be only *one* choice that ensures stability). Accordingly, we will guarantee that the choice of matching these two men to their best partners translates to *not* choosing the vertex they represent into the clique, and the other choice translates to choosing this vertex into the clique.

Now, having just the set M_V, we can encode selection of vertices into the clique, but we cannot ensure that the vertices we select indeed form a clique. For this purpose, we also have the set M_E which, in a manner similar to M_V, encodes selection of edges into the clique. By designing the instance in a way that the situation of the men in the man-optimal stable matching is significantly worse than that of the women in the woman-optimal stable matching, we are able to ensure that at most $2(k + k(k-1)/2)$ men in $M_V \cup M_E$ will not be assigned their best partners (here, we exploit the condition that $\mathsf{balance}(\mu) \leq \widehat{k}$ for a solution μ). Here the man m^* plays a crucial role—by using dummy men and women (in the sets $\widetilde{M}$ and $\widetilde{W}$) that prefer each other over all other people, we ensure that the situation of m^* is always “extremely bad” (from his viewpoint), while the situation of his partner, w^*, is always “excellent” (from her viewpoint).

At this point, we first need to ensure that the edges that we select indeed connect the vertices that we select. For this purpose, we carefully design our reduction so that when a pair of men representing some edge e obtain partners worse than those they have in the man-optimal stable matching, it must be that the men representing the endpoints of e have also obtained partners worse than those they have in the man-optimal stable matching, else stability will not be preserved—the partners of the men representing the endpoints of e will form blocking pairs together with the men representing e.

Finally we observe that we still need to ensure that among our $2(k + k(k-1)/2)$ distinguished men in $M_V \cup M_E$, which are associated with $k + k(k-1)/2$ selected elements (vertices and edges), there will be exactly $2k$ distinguished men from M_V and exactly $k(k-1)$ distinguished men from M_E, which would mean we have chosen k vertices and $k(k-1)/2$ edges. For this purpose, we construct an instance where for the women, it is only somewhat “beneficial” that the men in M_V will not be matched to their best partners, but it is extremely beneficial that the men in M_E will not be matched to their best partners. This objective is achieved by carefully placing dummy men (from $\widetilde{M}$) in the preference lists of women in W_E. By again exploiting the condition that $\mathsf{balance}(\mu) \leq \widehat{k}$ for a solution μ, we are able to ensure that there would be at least $k(k-1)$ distinguished men from M_E. We are now ready to present the formal details.

3.1. Formal description of the reduction

Next, we proceed with the formal presentation of our reduction by defining preference of every man $m \in M$, and thus constructing $\mathcal{L}_M$, given in Table 1. Accordingly, we define preference of every woman $w \in W$, and thus construct $\mathcal{L}_W$, given in Table 2.

Table 2
Preference list for every woman $w \in W$, and thus constructing $\mathcal{L}_W$. For any set X, $[X]$ denotes a fixed permutation of X. Here, $\deg_G(v)$ denotes the degree of the vertex v in G.

W	Preference list	
w_v^1	$m_v^2, [\{m_e^1 \in M_E : v \in e\}], [\{\widetilde{m}^i \in \widetilde{M} : i \le \vert E\vert - \deg_G(v)\}], m_v^1$	for all $w_v^1 \in W_V$
w_v^2	$m_v^1, [\{m_e^2 \in M_E : v \in e\}], [\{\widetilde{m}^i \in \widetilde{M} : i \le \vert E\vert - \deg_G(v)\}], m_v^2$	for all $w_v^2 \in W_V$
w_e^1	$m_e^2, \widetilde{m}^1, \widetilde{m}^2, \dots, \widetilde{m}^{\vert V\vert\vert E\vert}, m_e^1$	for all $w_e^1 \in W_E$
w_e^2	$m_e^1, \widetilde{m}^1, \widetilde{m}^2, \dots, \widetilde{m}^{\vert V\vert\vert E\vert}, m_e^2$	for all $w_e^2 \in W_E$
$\widetilde{w}^i$	$\widetilde{m}^i, m^*, m_{v_1}^1, m_{v_2}^1 \dots, m_{v_{\vert V\vert}}^1, m_{v_1}^2, m_{v_2}^2 \dots, m_{v_{\vert V\vert}}^2$	for $\widetilde{w}^i \in \widetilde{W}$, $i \in \{1, 2\}$
$\widetilde{w}^i$	$\widetilde{m}^i, m^*$	for all $\widetilde{w}^i \in \widetilde{W}$ s.t $i > 2$
w^*	m^*	

Finally, we define $\widehat{k} = |M| + \delta + 6(k + \dfrac{k(k-1)}{2})$. It is clear that the entire construction (under the assumptions that $\delta \ge 0$ and $|V| > k + \dfrac{k(k-1)}{2}$) can be performed in polynomial time.

3.2. The parameter

Our current objective is to verify that t is indeed bounded by a function of k. For this purpose, we first observe that for all $i \in \{1, 2, \dots, \delta\}$, it holds that $p_{\widetilde{m}^i}(\widetilde{w}^i) = p_{\widetilde{w}^i}(\widetilde{m}^i) = 1$. Therefore, for all $\mu \in \mathrm{SM}(\widehat{\mathcal{I}})$ and $i \in \{1, 2, \dots, \delta\}$, we have that $\mu(\widetilde{m}^i) = \widetilde{w}^i$, else $(\widetilde{m}^i, \widetilde{w}^i)$ would have formed a blocking pair for μ in $\widehat{\mathcal{I}}$.

Observation 3.1. *For all $\mu \in SM(\widehat{\mathcal{I}})$ and $i \in \{1, 2, \dots, \delta\}$, it holds that $\mu(\widetilde{m}^i) = \widetilde{w}^i$.*

Now, note that $\mathcal{A}(m^*) = \widetilde{W} \cup \{w^*\}$. Thus, by Observation 3.1, we have that for all $\mu \in SM(\widehat{\mathcal{I}})$, either m^* is unmatched or $\mu(m^*) = w^*$. However, $\mathcal{A}(w^*) = \{m^*\}$, which implies that in the former case, (m^*, w^*) forms a blocking pair. Thus, we also have the following observation.

Observation 3.2. *For all $\mu \in SM(\widehat{\mathcal{I}})$, it holds that $\mu(m^*) = w^*$.*

Let us proceed by identifying the man-optimal μ_M and the woman-optimal μ_W stable matchings. For this purpose, we first define a matching μ'_M as follows.

- For all $m_v^i \in M_V$: $\mu'_M(m_v^i) = w_v^i$.
- For all $m_e^i \in M_E$: $\mu'_M(m_e^i) = w_e^i$.
- For all $\widetilde{m}^i \in \widetilde{M}$: $\mu'_M(\widetilde{m}^i) = \widetilde{w}^i$.
- $\mu'_M(m^*) = w^*$.

Lemma 3.2. *It holds that $\mu_M = \mu'_M$.*

Proof. Since for all $i \in \{1, 2, \dots, \delta\}$, it holds that $\mu'_M(\widetilde{m}^i) = \widetilde{w}^i$ and $p_{\widetilde{m}^i}(\widetilde{w}^i) = p_{\widetilde{w}^i}(\widetilde{m}^i) = 1$, we have that there cannot exist a blocking pair with at least one person from $\widetilde{M} \cup \widetilde{W}$. Now, notice that for every $m \in M$, including m^*, the woman most preferred by m who is outside $\widetilde{W}$ is also the one with whom it is matched. Therefore, there cannot exist any blocking pair for μ'_M, and by Observation 3.1, we further conclude that indeed $\mu_M = \mu'_M$. $\square$

Now, we define a matching μ'_W as follows.

- For all $w_v^1 \in W_V$: $\mu'_W(w_v^1) = m_v^2$, and for all $w_v^2 \in W_V$: $\mu'_W(w_v^2) = m_v^1$.
- For all $w_e^1 \in W_E$: $\mu'_W(w_e^1) = m_e^2$, and for all $w_e^2 \in W_E$: $\mu'_W(w_e^2) = m_e^1$.
- For all $\widetilde{w}^i \in \widetilde{W}$: $\mu'_W(\widetilde{w}^i) = \widetilde{m}^i$.
- $\mu'_W(w^*) = m^*$.

Lemma 3.3. *It holds that $\mu_W = \mu'_W$.*

Proof. In the matching μ'_W, every woman is matched with the man she prefers the most. Thus, it is immediate that $\mu_W = \mu'_W$. $\square$

As a corollary to Lemmata 3.2 and 3.3, we obtain the following result.

Corollary 3.1. *$O_M = |M| + \delta$ and $O_W = |W|$.*

Proof. First, note that

$$\begin{aligned} O_M &= \sum_{(m,w)\in\mu_M} p_m(w) \\ &= p_{m^*}(\mu_M(m^*)) + \sum_{m\in M\setminus\{m^*\}} p_m(\mu_M(m)) = (\delta+1) + (|M|-1) = |M| + \delta. \end{aligned}$$

Second, note that

$$O_W = \sum_{(m,w)\in\mu_W} p_w(m) = \sum_{w\in W} p_w(\mu_W(w)) = |W|. \quad \square$$

We are now ready to bound t.

Lemma 3.4. *The parameter t associated with $\widehat{\mathcal{I}}$ is equal to $6(k + \frac{k(k-1)}{2})$.*

Proof. By the definition of t, we have that

$$\begin{aligned} t &= \widehat{k} - \max\{O_M, O_W\} \\ &= |M| + \delta + 6(k + \frac{k(k-1)}{2}) - \max\{|M| + \delta, |W|\} = 6(k + \frac{k(k-1)}{2}). \quad \square \end{aligned}$$

3.3. Correctness

First, from Lemma 3.2 and Proposition 2.2 we derive the following useful observation.

Observation 3.3. *Every $\mu \in SM(\widehat{\mathcal{I}})$ matches all people in $M \cup W$.*

Next, we proceed to state the first direction necessary to conclude that the input instance $\mathcal{I}$ of Clique and our instance $\widehat{\mathcal{I}}$ of Above-Max BSM are equivalent.

Lemma 3.5. *If $\mathcal{I}$ is a* Yes*-instance, then $\widehat{\mathcal{I}}$ is a* Yes*-instance.*

Proof. Suppose that $\mathcal{I}$ is a Yes-instance, and let U be the vertex set of a clique on k vertices in G. We denote $M_V^U = \{m_v^i \in M_V : v \in U\}$ and $M_E^U = \{m_{\{u,v\}}^i \in M_E : u, v \in U\}$. Then, we define a matching μ as follows.

- For all $m_v^i \in M_V$:
  - If $m_v^i \in M_V^U$: $\mu(m_v^i) = w_v^{3-i}$.
  - Else: $\mu(m_v^i) = w_v^i$.
- For all $m_e^i \in M_E$:
  - If $m_e^i \in M_E^U$: $\mu(m_e^i) = w_e^{3-i}$.
  - Else: $\mu(m_e^i) = w_e^i$.
- For all $\widetilde{m}^i \in \widetilde{M}$: $\mu(\widetilde{m}^i) = \widetilde{w}^i$.
- $\mu(m^*) = w^*$.

We claim that $\mu \in SM(\widehat{\mathcal{I}})$ and $\mathsf{balance}(\mu) \le \widehat{k}$, which would imply that $\widehat{\mathcal{I}}$ is a Yes-instance. To this end, we first show that $\mu \in SM(\widehat{\mathcal{I}})$. Since for all $i \in \{1, 2, \ldots, \delta\}$, it holds that $\mu(\widetilde{m}^i) = \widetilde{w}^i$ and $p_{\widetilde{m}^i}(\widetilde{w}^i) = p_{\widetilde{w}^i}(\widetilde{m}^i) = 1$, we have that there cannot exist a blocking pair with at least one person from $\widetilde{M} \cup \widetilde{W}$. Thus, there can also not be a blocking pair with any person from $\{m^*, w^*\}$.

On the one hand, notice that for every $m \in (M_V \setminus M_V^U) \cup (M_E \setminus M_E^U) \cup \{m^*\}$, the woman most preferred by m who is outside $\widetilde{W}$ is also the one with whom it is matched. Thus, no man in $(M_V \setminus M_V^U) \cup (M_E \setminus M_E^U) \cup \{m^*\}$ can belong to a blocking pair. Moreover, the set of acceptable partners of any woman in W_E matched to a man in $M_E \setminus M_E^U$ is a subset of $\widetilde{M} \cup (M_E \setminus M_E^U)$, and therefore such a woman cannot belong to a blocking pair. On the other hand, let W' denote the set of every woman that is matched to a man $m \in M_V^U \cup M_E^U$. Then, for every $w \in W'$, the man most preferred by w is also the one with whom she is matched. Therefore, no woman in W' can belong to a blocking pair. Hence, we also conclude that no woman in W_E can belong to a blocking pair.

Thus, if there exists a blocking pair, it must consist of a man $m \in M_V^U \cup M_E^U$ and a woman $w \in W_V \setminus W'$. Suppose, by way of contradiction, that there exists such a blocking pair (m, w). First, let us assume that $m = m_v^i \in M_V^U$. In this case, since apart from w_v^i, all women in $\mathcal{A}(m_v^i)$ belong to $\widetilde{W} \cup \{\mu(m_v^i)\}$, we deduce that $w = w_v^i$. However, w_v^i prefers $\mu(w_v^i)$ over m_v^i, and thus we reach a contradiction. Next, we assume that $m = m_{\{u,v\}}^i \in M_E^U$. In this case, it must hold that w is either w_v^i or w_u^i. Without loss of generality, we assume that $w = w_v^i$. However, since $m_{\{u,v\}}^i \in M_E^U$, we have that $v \in U$. Therefore, $\mu(w_v^i) = m_v^{3-i}$. Since w_v^i prefers m_v^{3-i} over $m_{\{u,v\}}^i$, we reach a contradiction.

It remains to prove that $\mathsf{balance}(\mu) \le \widehat{k}$. To this end, we need to show that

$$\max\{\sum_{(m,w)\in\mu} p_m(w), \sum_{(m,w)\in\mu} p_w(m)\} \le |M| + \delta + 6(k + \frac{k(k-1)}{2}).$$

First, note that

$$\begin{aligned}\sum_{(m,w)\in\mu} p_m(w) &= \sum_{m\in M_V^U} p_m(\mu(m)) + \sum_{m\in M_V\setminus M_V^U} p_m(\mu(m)) + \sum_{m\in M_E^U} p_m(\mu(m)) \\ &\quad + \sum_{m\in M_E\setminus M_E^U} p_m(\mu(m)) + \sum_{m\in\widetilde{M}} p_m(\mu(m)) + p_{m^*}(\mu(m^*)) \\ &= 4|M_V^U| + |M_V \setminus M_V^U| + 4|M_E^U| + |M_E \setminus M_E^U| + |\widetilde{M}| + \delta + 1 \\ &= |M| + \delta + 3(|M_V^U| + |M_E^U|) = |M| + \delta + 6(k + \frac{k(k-1)}{2}).\end{aligned}$$

Second, note that

$$\begin{aligned}\sum_{(m,w)\in\mu} p_w(m) &= \sum_{m\in M_V^U} p_{\mu(m)}(m) + \sum_{m\in M_V\setminus M_V^U} p_{\mu(m)}(m) + \sum_{m\in M_E^U} p_{\mu(m)}(m) \\ &\quad + \sum_{m\in M_E\setminus M_E^U} p_{\mu(m)}(m) + \sum_{m\in\widetilde{M}} p_{\mu(m)}(m) + p_{\mu(m^*)}(m^*) \\ &= |M_V^U| + |M_V \setminus M_V^U|(|E|+2) + |M_E^U| + |M_E \setminus M_E^U|(|V||E|+2) + |\widetilde{M}| + 1 \\ &= |M| + 2(|V|-k)(|E|+1) + 2(|E| - \frac{k(k-1)}{2})(|V||E|+1) \\ &= |M| + \delta + 6(k + \frac{k(k-1)}{2}).\end{aligned}$$

This concludes the proof of the lemma. □

We now turn to prove the second direction.

Lemma 3.6. *If $\widehat{\mathcal{I}}$ is a* Yes*-instance, then $\mathcal{I}$ is a* Yes*-instance.*

Proof. Suppose that $\widehat{\mathcal{I}}$ is a Yes-instance, and let μ be a stable matching such that $\mathsf{balance}(\mu) \le \widehat{k}$. By Observations 3.1 and 3.2, it holds that

- For all $i \in \{1, 2, \ldots, \delta\}$: $\mu(\widetilde{m}^i) = \widetilde{w}^i$.
- $\mu(m^*) = w^*$.

Thus, since Observation 3.3 implies that all vertices in M_V should be matched by μ, we deduce that

- For all $v \in V$: Either both $\mu(m_v^1) = w_v^1$ and $\mu(m_v^2) = w_v^2$ or both $\mu(m_v^2) = w_v^1$ and $\mu(m_v^1) = w_v^2$.

Let U denote the set of every $v \in V$ such that $\mu(m_v^2) = w_v^1$ and $\mu(m_v^1) = w_v^2$. Moreover, denote $M_V^U = \{m_v^i \in M_V : v \in U\}$. By the item above, and since all vertices in M_E should also be matched by μ, we further deduce that

- For all $e \in E$: Either both $\mu(m_e^1) = w_e^1$ and $\mu(m_e^2) = w_e^2$ or both $\mu(m_e^2) = w_e^1$ and $\mu(m_e^1) = w_e^2$.

Let S denote the set of every $e \in E$ such that $\mu(m_e^2) = w_e^1$ and $\mu(m_e^1) = w_e^2$. Moreover, denote $M_E^S = \{m_e^i \in M_E : e \in S\}$. If there existed $\{u, v\} \in S$ such that $u \notin U$, then $(m_{\{u,v\}}^1, w_u^1)$ would have formed a blocking pair, which contradicts the fact that μ is a stable matching. Thus, we have that the set of endpoints of the edges in S is a subset of U.

We claim that $|U| = k$ and that U is the vertex set of a clique in G, which would imply that $\mathcal{I}$ is a YES-instance. Since we have argued that the set of endpoints of the edges in S is a subset of U, it is sufficient to show that $|U| \leq k$ and $|S| \geq \frac{k(k-1)}{2}$ (note that $|S| \geq \frac{k(k-1)}{2}$ implies that $|U| \geq k$), as this would imply that U is indeed the vertex set of a clique on k vertices in G. First, since $\mathsf{balance}(\mu) \leq \widehat{k}$, we have that $\sum_{(m,w)\in\mu} p_m(w) \leq |M| + \delta + 6(k + \frac{k(k-1)}{2})$. Now, note that

$$\begin{aligned}\sum_{(m,w)\in\mu} p_m(w) &= \sum_{m\in M_V^U} p_m(\mu(m)) + \sum_{m\in M_V\setminus M_V^U} p_m(\mu(m)) + \sum_{m\in M_E^S} p_m(\mu(m))\\ &\quad+ \sum_{m\in M_E\setminus M_E^S} p_m(\mu(m)) + \sum_{m\in\widetilde{M}} p_m(\mu(m)) + p_{m^*}(\mu(m^*))\\ &= 4|M_V^U| + |M_V \setminus M_V^U| + 4|M_E^S| + |M_E \setminus M_E^S| + |\widetilde{M}| + \delta + 1\\ &= |M| + \delta + 6(|U| + |S|).\end{aligned}$$

Thus, we deduce that $|U| + |S| \leq k + \frac{k(k-1)}{2}$. Now, observe that since $\mathsf{balance}(\mu) \leq \widehat{k}$, we also have that $\sum_{(m,w)\in\mu} p_w(m) \leq |M| + \delta + 6(k + \frac{k(k-1)}{2})$. Here, on the one hand we note that

$$\begin{aligned}\sum_{(m,w)\in\mu} p_w(m) &= \sum_{m\in M_V^U} p_{\mu(m)}(m) + \sum_{m\in M_V\setminus M_V^U} p_{\mu(m)}(m) + \sum_{m\in M_E^S} p_{\mu(m)}(m)\\ &\quad+ \sum_{m\in M_E\setminus M_E^S} p_{\mu(m)}(m) + \sum_{m\in\widetilde{M}} p_{\mu(m)}(m) + p_{\mu(m^*)}(m^*)\\ &= |M_V^U| + |M_V \setminus M_V^U|(|E|+2) + |M_E^S|\\ &\quad+ |M_E \setminus M_E^S|(|V||E|+2) + |\widetilde{M}| + 1\\ &= |M| + 2(|V| - |U|)(|E|+1) + 2(|E| - |S|)(|V||E|+1)\\ &= |M| + 2(|V| + |E| + |V||E| + |V||E|^2)\\ &\quad -2|U|(|E|+1) - 2|S|(|V||E|+1).\end{aligned}$$

On the other hand, we note that

$$\begin{aligned}\widehat{k} &= |M| + \delta + 6(k + \frac{k(k-1)}{2})\\ &= |M| + 2(|V| + |E| + |V||E| + |V||E|^2) - k(4 + 4k + 2|E| + (k-1)|V||E|) + 6(k + \frac{k(k-1)}{2})\\ &= |M| + 2(|V| + |E| + |V||E| + |V||E|^2) - 2k(|E|+1) - k(k-1)(|V||E|+1).\end{aligned}$$

Thus, we have that

$$|U|(|E|+1) + |S|(|V||E|+1) \geq k(|E|+1) + \frac{k(k-1)}{2}(|V||E|+1)$$

Recall that we have also shown that $|U| + |S| \leq k + \frac{k(k-1)}{2}$. Thus, since $|U| \leq k + \frac{k(k-1)}{2} - |S|$, to satisfy the above equation it must hold that $|S| \geq \frac{k(k-1)}{2}$. Since $|U| + |S| \leq k + \frac{k(k-1)}{2}$, we deduce that $|U| \leq k$. This, as we have argued earlier, finished the proof. □

This concludes the proof of Theorem 1.

4. Kernel

In this section, we design a kernelization algorithm for ABOVE-MIN BSM. More precisely, we prove Theorem 2.

4.1. *Functional balanced stable marriage*

To prove Theorem 2, we first prove the following result for the ABOVE-MIN FBSM problem.

Lemma 4.1. ABOVE-MIN FBSM *admits a kernel with at most* $2t$ *men, at most* $2t$ *women, and such that the image of the preference function of each person is a subset of* $\{1, 2, \ldots, t+1\}$. *The kernel can be found in time* $\mathcal{O}(\ell n^2)$, *where* ℓ *denotes the number of acceptable pairs and n denotes the number of agents.*

To obtain the desired kernelization algorithm, we execute the following steps if $t \geq 0$. (If $t < 0$, then $k < \min\{O_M, O_W\} \leq \max\{O_M, O_W\}$. We prove in Lemma 4.2 that the given instance is a No-instance.)

1. **Cleaning Prefixes and Suffixes.** Simplify the preference functions by "cleaning" suffixes and thereby also "cleaning" prefixes.
   In this step, we remove from a person a's set of acceptable partners, $\mathcal{A}(a)$, an individual who will never be matched to the former in *any* stable matching of the given instance. Intuitively speaking, we remove people who are either in the *prefix* or the *suffix* of a's preference list. A prefix for a man (woman) consists of women (men) who are ranked better than his (her) partner in the man- (woman-) optimal stable matching. Analogously, a suffix for a man (woman) consists of women (men) who are ranked lower than his (her) partner in the woman- (man-) optimal stable matching. For a man m, we execute this step by removing from $\mathcal{A}(m)$ all those women whose f_m–value is either lower than $f_m(\mu_M(m))$, the function value of m's man-optimal stable matching partner, or higher than $f_m(\mu_W(m))$, the function value of m's woman-optimal stable matching partner. Symmetrically, for a woman w, we reduce $\mathcal{A}(w)$ by removing all those men whose f_w–value is lower than $f_w(\mu_W(w))$ or higher than $f_w(\mu_M(w))$. Since the domain of a preference function f_a is the set of acceptable partners $\mathcal{A}(a)$, this step restricts the preference function itself.
2. **Perfect Matching.** Zoom into the set of people matched by every stable matching.
   When the preferences are complete, that is, each person is an acceptable partner for every individual on the other side, then the stable matchings are perfect. Otherwise, we know that by the Rural Hospital theorem [45], if a person (man/woman) is unmatched in the man-optimal stable matching, then that person will be unmatched in every stable matching in the instance. Consequently, we can remove such agents from the instance and restrict the preference functions appropriately. This ensures that in the resulting instance, every stable matching is a perfect matching.
3. **Overcoming Sadness.** Bound the number of "sad" people. A person is sad if (s)he is not a "happy" person, defined to be one who has the same partner in every stable matching. Thus, an individual a is sad if a's best stable matching partner is not also a's worst stable matching partner.
   In this step, we give a bound on the number of sad people. Towards this, we check the progress we have made so far in reducing the size of the instance, namely by reducing the number of men and the number of women. We show that if the first two steps are not applicable, then the number of sad people in the instance convey useful properties about the instance. Specifically, if the number of sad women/men in the instance is more than $2t$, then it is a No-instance. Conversely, if there are no sad men (or women) in the instance, then it is a Yes-instance if and only if the balance()-value of the man-optimal (and woman-optimal) stable matching is at most k.
4. **Marrying Happy People.** Remove "happy" people from the instance.
   Note that a happy man is always matched to the same happy woman in every stable matching; such a pair is better known as a *fixed pair* in the stable matching literature. By removing a fixed pair, say (m_h, w_h), from the instance, we are creating a new instance in which the set of stable matchings are in bijective correspondence with the set of stable matchings in the original instance. The balance values of each matchings, however, will go down by a fixed and known quantity, $f_{m_h}(w_h)$ or $f_{w_h}(m_h)$. So when we remove a fixed pair from our instance, we transfer its contribution to the balance()-value of the original instance to the new instance. This ensures that the balance value of each of the matchings in the new instance is the same as the value of the corresponding matching in the original instance.
5. **Removing High Value Partners.** Obtain "compact" preference functions by truncating "high-values".
   For each person, the goal is to remove acceptable partners who if matched to (in a stable matching) will inexorably raise the balance value of the resulting matching to higher than k, irrespective of the other matching pairs. Thus, it follows that such a matching cannot be a yes certificate (the reason the instance is a Yes-instance) for our instance. Moreover, if it is a No-instance, then the balance value of every stable matching must be higher than k. Hence, removing a very high value pair from the instance (as long as there are others in the instance[6]) will not destroy the equivalence property of the resulting instance.
   Specifically, we know that if for a man m and woman w, who are each other acceptable partner, we either have $f_m(w) > f_m(\mu_M(m)) + (k - O_M)$ or $f_w(m) > f_w(\mu_W(w)) + (k - O_W)$, then any stable matching that contains the pair (m, w) will have balance()-value greater than k. Hence, it is safe to remove w from the set $\mathcal{A}(m)$ and m from $\mathcal{A}(w)$.
   Since we cleaned the prefixes, it may be that for some man m and woman w, the pre-images $f_m^{-1}(1)$ and $f_w^{-1}(1)$ may not be m's man-optimal stable matching partner and w's woman-optimal stable matching partner, respectively. Recall that we will apply the steps repeatedly as long as they are applicable. Hence, in this scenario we perform the following step.
6. **Shrinking Gaps.** Shrink some of the gaps created by previous steps.
   We perform this step by decreasing $f_m(w')$ by 1 for every woman $w' \in \mathcal{A}(m)$. Similarly, for every $m' \in \mathcal{A}(w)$, we decrease $f_w(m')$ by 1. Since, we are reducing the preference function value of m and w's best stable matching partner, and by Step 1 we know that both of them will be matched in every stable matching, it is necessary that we reduce the value of k (the target balance()-value) by 1.
   This completes the construction of the kernel for Above-Min FBSM.

[6] Since our instance only contains sad men and women, it follows readily that there are more than one stable matchings.

One way to see the intuition behind the safeness of each of our reduction rules is to see their affect on the stable matching lattice of the original instance.

Affect on the stable matching lattice: Let $\mathcal{I}$ denote the initial instance. The applications of Steps 1–3 leave the stable matching lattice of $\mathcal{I}$ intact. Step 4 creates an instance, denoted by $\mathcal{J}$, whose stable matching lattice consists of submatchings of the original instance such that the stable matchings in the two instances are in bijective correspondence. Specifically, each stable matching μ in $\mathcal{I}$ gives rises to a stable matching μ' in $\mathcal{J}$ such that $\mu \setminus \mu'$ is the set of fixed pairs in $\mathcal{I}$. Step 5 actually results in a lattice that is a subset of the lattice of the initial instance. Finally it is easy to see that Step 6 does not remove or alter partnerships and it does not have any affect on the stable matching lattice of the initial instance.

Obtaining a kernel for Above-Min BSM: Using the kernel for Above-Min FBSM, we obtain a kernel for Above-Min BSM by interpreting the preference functions as preference lists. The main roadblock towards this is that the preference functions may have "gaps" in the function values, i.e. the function f_m (f_w) for some man m (woman w) may not be surjective in the range $[f_m(\mu_M(m)), f_m(\mu_W(m))]$ ($[f_w(\mu_W(w)), f_w(\mu_M(w))]$). Notably, while creating the preference list for m, we cannot have gaps in the preference list of m.

Maximum function value is bounded. Since Step 5 is not applicable, for any pair (m, w) who are each other's acceptable partner, we have

$$f_m(w) \leq f_m(\mu_M(m)) + (k - O_M) \leq f_m(\mu_M(m)) + k - \min\{O_M, O_W\} = f_m(\mu_M(m)) + t. \quad \text{Symmetrically, we have } f_w(m) \leq f_w(\mu_W(m)) + t. \tag{1}$$

Since Step 6 is not applicable, it follows that $f_m(\mu_M(m)) = 1$ and $f_w(\mu_W(w)) = 1$. Hence, we have ensured that $t + 1$ is the maximum value any preference function can take.

7. **Filling the "gaps" in the preference function.** For a man m, consider its preference function values in increasing order: $f_m(x_1), f_m(x_2), \ldots . f_m(x_{t'})$, where $t' \leq t + 1$, due to Equation (1).
   If any of these two adjacent values are not actually consecutive, then we have a "gap" in our preference function. For some $i \in [t']$ let $s_i = f_m(x_{i+1}) - f_m(x_i) > 1$, then we fill the "gap" between $f_m(x_i)$ and $f_m(x_{i+1})$ by $s_i - 1$ dummy women $y_1, \ldots y_{s_i - 1}$, where $s_i \leq t' - 1 \leq t$. We do this for every gap that exists in the preference function of m. Since the maximum value of $f_m(\cdot)$ is at most $t + 1$, thus we require no more than $t - 1$ dummy women to fill the gaps in m's preference function. Note that the same set of $t - 1$ dummy women can be used to fill the gaps in each man's preference function. Additionally, we use $t - 1$ dummy men to create dummy happy pairs for these women. Formally, this is done by defining another preference function g_m which extends the domain of f_m such that there are no gaps in g_m. The preference function g_m can be interpreted as a preference list in a straightforward fashion. Doing this for every man/woman whose preference function has "gaps" leads to a kernel for Above-Min BSM. Thus, overall the kernel for Above-Min BSM has t additional men and t additional women.

In our upcoming discussion, we elaborate on each of these reduction rules formally. In what follows, we let $\mathcal{I}$ denote our current instance of Above-Min FBSM. Initially, this instance is the input instance, but as the execution of our algorithm progresses, the instance is modified. The reduction rules that we present are applied *exhaustively* in the order of their presentation. In other words, at each point of time, the first rule whose condition is true is the one that we apply next. In particular, the execution terminates once the value of t drops below 0, as implied by the next rule.

It is worth mentioning that the phenomenon captured by some of our initial Reduction Rules such as 4.2 and 4.3 have previously been studied (see [22] Theorem 1.2.5 and 1.4.2), albeit in a different context or language. The safeness of those rules, therefore, may well be a foregone conclusion for some readers. However, we present all the proofs here for the sake of completeness. From Proposition 2.1 we can infer that each of our reduction rules can indeed be implemented in time polynomial in the number of men and women.

Reduction Rule 4.1. *If $k < \max\{O_M, O_W\}$, then return* No.

Lemma 4.2. *Reduction Rule 4.1 is safe.*

Proof. For every $\mu \in \text{SM}$, it holds that $\mathsf{balance}(\mu) \geq \max\{O_M, O_W\}$. Thus, if $k < \max\{O_M, O_W\}$, then every $\mu \in \text{SM}$ satisfies $\mathsf{balance}(\mu) > k$. In this case, we conclude that $\mathsf{Bal} > k$, and therefore $\mathcal{I}$ is a No-instance. $\square$

Note that if $t < 0$, then $k < \min\{O_M, O_W\} \leq \max\{O_M, O_W\}$.

Cleaning prefixes and suffixes. We begin by modifying the images of the preference functions. We remark that it is *necessary* to perform this step first, after Reduction Rule 4.1, otherwise, the following steps would not be correct. To clean prefixes while ensuring *both* safeness and that the parameter t does not increase, we would actually need to clean *suffixes* first. Formally, we define suffixes as follows.

Definition 4.1. *Let (m, w) denote an acceptable pair. If m is matched by μ_W and $f_m(w) > f_m(\mu_W(m))$, then we say that w* belongs to the suffix of *m. Similarly, if w is matched by μ_M and $f_w(m) > f_w(\mu_M(w))$, then we say that m belongs to the suffix of w.*

By Proposition 2.3, we have the following observation.

Observation 4.1. *Let (m, w) denote an acceptable pair such that one of its members belongs to the suffix of the other member. Then, there is no $\mu \in \mathrm{SM}(\mathcal{I})$ that matches m with w.*

For every person a, let $\mathsf{worst}(a)$ be the person in $\mathcal{A}(a)$ to whom f_a assigns its worst preference value. More precisely, $\mathsf{worst}(a) = \mathrm{argmax}_{b\in\mathcal{A}(a)} f_a(b)$. We will now clean suffixes.

Reduction Rule 4.2. *If there exists a person a such that $\mathsf{worst}(a)$ belongs to the suffix of a, then define the preference functions as follows.*

- $f'_a = f_a|_{\mathcal{A}(a)\setminus\{\mathsf{worst}(a)\}}$ *and* $f'_{\mathsf{worst}(a)} = f_{\mathsf{worst}(a)}|_{\mathcal{A}(\mathsf{worst}(a))\setminus\{a\}}$.
- *For all $b \in M \cup W \setminus \{a, \mathsf{worst}(a)\}$: $f'_b = f_b$*

The new instance is $\mathcal{J} = (M, W, \{f'_{m'}\}_{m'\in M}, \{f'_{w'}\}_{w'\in W}, k)$.

Lemma 4.3. *Reduction Rule 4.2 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Proof. By the definition of the new preference functions, we have that for every $\mu \in \mathrm{SM}(\mathcal{I}) \cap \mathrm{SM}(\mathcal{J})$, it holds that $\sum_{(m,w)\in\mu} f_m(w) = \sum_{(m,w)\in\mu} f'_m(w)$ and $\sum_{(m,w)\in\mu} f_w(m) = \sum_{(m,w)\in\mu} f'_w(m)$. In particular, this means that to conclude that $\mathsf{Bal}(\mathcal{I}) = \mathsf{Bal}(\mathcal{J})$ (which implies safeness) as well as that $O_M(\mathcal{I}) = O_M(\mathcal{J})$ and $O_W(\mathcal{I}) = O_W(\mathcal{J})$ (which implies that $t(\mathcal{I}) = t(\mathcal{J})$), it is sufficient to show that $\mathrm{SM}(\mathcal{I}) = \mathrm{SM}(\mathcal{J})$. For this purpose, first consider some $\mu \in \mathrm{SM}(\mathcal{I})$. By Observation 4.1, it holds that $(a, \mathsf{worst}(a)) \notin \mu$. Hence, μ is a matching in $\mathcal{J}$. Moreover, if μ has a blocking pair in $\mathcal{J}$, then by the definition of the new preference functions, it is also a blocking pair in $\mathcal{I}$. Since μ is stable in $\mathcal{I}$, we have that $\mu \in \mathrm{SM}(\mathcal{J})$.

In the second direction, consider some $\mu \in \mathrm{SM}(\mathcal{J})$. Then, it is clear that μ is a matching in $\mathcal{I}$. Moreover, if μ has a blocking pair (m, w) in $\mathcal{I}$ that is not $(a, \mathsf{worst}(a))$, then (m, w) is an acceptable pair in $\mathcal{J}$, and therefore by the definition of the new preference functions, we have that (m, w) is also a blocking pair in $\mathcal{J}$. Hence, since μ is stable in $\mathcal{J}$, the only pair that can block μ in $\mathcal{I}$ is $(a, \mathsf{worst}(a))$. Thus, to show that $\mu \in \mathrm{SM}(\mathcal{I})$, it remains to prove that $(a, \mathsf{worst}(a))$ cannot block μ in $\mathcal{I}$. Suppose, by way of contradiction, that $(a, \mathsf{worst}(a))$ blocks μ in $\mathcal{I}$. In particular, this means that $f_a(\mathsf{worst}(a)) < f_a(\mu(a))$. However, this contradicts the definition of $\mathsf{worst}(a)$. $\square$

By cleaning suffixes, we actually also accomplish the objective of cleaning prefixes, which are defined as follows.

Definition 4.2. *Let (m, w) denote an acceptable pair. If m is matched by μ_M and $f_m(w) < f_m(\mu_M(m))$, then we say that w* belongs to the prefix of *m. Similarly, if w is matched by μ_W and $f_w(m) < f_w(\mu_W(w))$, then we say that m belongs to the prefix of w.*

Lemma 4.4. *Let $\mathcal{I}$ be an instance of* Above-Min FBSM *on which Reduction Rules 4.1 to 4.2 have been exhaustively applied. Then, there does not exist an acceptable pair (m, w) such that one of its members belongs to the prefix of the other one.*

Proof. Suppose, by way of contradiction, that there exists an acceptable pair (m, w) such that one of its members belongs to the prefix of the other one. Without loss of generality, assume that w belongs to the prefix of m. Then, $f_m(w) < f_m(\mu_M(m))$. Since μ_M is a stable matching, it cannot be blocked by (m, w), which means that w is matched by μ_M and $f_w(\mu_M(w)) < f_w(m)$. Thus, we have that m belongs to the suffix of w, which contradicts the assumption that Reduction Rule 4.2 was applied exhaustively. $\square$

Corollary 4.1. *Let $\mathcal{I}$ be an instance of* Above-Min FBSM *on which Reduction Rules 4.1 to 4.2 have been exhaustively applied. Then, for every acceptable pair (m, w) in $\mathcal{I}$ where m and w are matched (not necessarily to each other) by both μ_M and μ_W, it holds that $f_m(\mu_M(m)) \leq f_m(w) \leq f_m(\mu_W(m))$ and $f_w(\mu_W(w)) \leq f_w(m) \leq f_w(\mu_M(w))$.*

Perfect matching. Having Corollary 4.1 at hand, we are able to provide a simple rule that allows us to assume that every solution matches all people.

Reduction Rule 4.3. *If there exists a person unmatched by μ_M, then let M' and W' denote the subsets of men and women, respectively, who are matched by μ_M. For each $a \in M' \cup W'$, denote $\mathcal{A}'(a) = \mathcal{A}(a) \cap (M' \cup W')$, and define $f'_a = f_a|_{\mathcal{A}'(a)}$. The new instance is $\mathcal{J} = (M', W', \{f'_m\}_{m\in M'}, \{f'_w\}_{w\in W'}, k)$.*

Note that if a person is unmatched in μ_M, then, due to Proposition 2.2, (s)he is unmatched in every stable matching in $\mathcal{I}$. To prove the safeness of this rule, we first prove the following lemma.

Lemma 4.5. *Let $\mathcal{I}$ be an instance of* Above-Min FBSM *on which Reduction Rules 4.1 to 4.2 have been exhaustively applied. Then, for every person a not matched by μ_M, it holds that $\mathcal{A}(a) = \emptyset$.*

Proof. Let a be a person not matched by μ_M. Then, by Proposition 2.2, it holds that a is not matched by any stable matching. We assume w.l.o.g. that a is a man, say m. First, note that $\mathcal{A}(m)$ cannot contain any woman w that is not matched by some stable matching, else (m, w) would have formed a blocking pair for that stable matching. Second, we claim that $\mathcal{A}(m)$ cannot contain a woman w that is matched by some stable matching. Suppose, by way of contradiction, that this claim is false. Then, by Proposition 2.2, it holds that $\mathcal{A}(m)$ contains a woman w that is matched by μ_M. We have that w prefers $\mu_M(w)$ over m, else (m, w) would have formed a blocking pair for μ_M, which is impossible as μ_M is a stable matching. However, this implies that m belongs to the suffix of w, which contradicts the supposition that Reduction Rule 4.2 has been exhaustively applied. We thus conclude that $\mathcal{A}(a) = \emptyset$. □

Lemma 4.6. *Reduction Rule 4.3 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Proof. By the definition of the new preference functions, we have that for every $\mu \in \mathrm{SM}(\mathcal{I}) \cap \mathrm{SM}(\mathcal{J})$, it holds that $\sum_{(m,w)\in\mu} f_m(w) = \sum_{(m,w)\in\mu} f'_m(w)$ and $\sum_{(m,w)\in\mu} f_w(m) = \sum_{(m,w)\in\mu} f'_w(m)$. To conclude that the lemma is correct, it is thus sufficient to argue that $\mathrm{SM}(\mathcal{I}) = \mathrm{SM}(\mathcal{J})$. For this purpose, first consider some $\mu \in \mathrm{SM}(\mathcal{I})$. By Proposition 2.2, we have that μ is also a matching in $\mathcal{J}$.

Moreover, if μ has a blocking pair in $\mathcal{J}$, then by the definition of the new preference functions, it is also a blocking pair in $\mathcal{I}$. Since μ is stable in $\mathcal{I}$, we have that $\mu \in \mathrm{SM}(\mathcal{J})$.

In the second direction, consider some $\mu \in \mathrm{SM}(\mathcal{J})$. Then, it is clear that μ is a matching in $\mathcal{I}$. By Lemma 4.5, if μ has a blocking pair (m, w) in $\mathcal{I}$, then both $m \in M'$ and $w \in W'$. However, for such a blocking pair (m, w), we have that (m, w) is an acceptable pair in $\mathcal{J}$, and therefore by the definition of the new preference functions, we have that (m, w) is also a blocking pair in $\mathcal{J}$. Hence, since μ is stable in $\mathcal{J}$, we conclude that μ is also stable in $\mathcal{I}$. □

By Proposition 2.2, from now onwards, we have that for the given instance, any stable matching is a perfect matching. Due to this observation, we can denote $n = |M| = |W|$, and for any stable matching μ, we have the following equalities.

$$\sum_{(m,w)\in\mu} f_m(w) = \sum_{m\in M} f_m(\mu(m)); \qquad \sum_{(m,w)\in\mu} f_w(m) = \sum_{w\in W} f_w(\mu(w)). \tag{I}$$

Overcoming sadness. As every stable matching is a perfect matching, every person is matched by every stable matching, including the man-optimal and woman-optimal stable matchings. Thus, it is well defined to classify the people who do not have the same partner in the man-optimal and woman-optimal stable matchings as “sad”. That is,

Definition 4.3. *A person $a \in M \cup W$ is* sad *if $\mu_M(a) \neq \mu_W(a)$.*

We let M_S and W_S denote the sets of sad men and sad women, respectively. People who are not sad are termed *happy*. Accordingly, we let M_H and W_H denote the sets of happy men and happy women, respectively. Note that $M_S = \emptyset$ if and only if $W_S = \emptyset$. Moreover, note that by the definition of μ_M and μ_W, for a happy person a it holds that a and $\mu_M(a) = \mu_W(a)$ are matched to one another by every stable matching. Let us now bound the number of sad people in a Yes-instance.

Reduction Rule 4.4. *If $|M_S| > 2t$ or $|W_S| > 2t$, then return* No.

Lemma 4.7. *Reduction Rule 4.4 is safe.*

Proof. We only prove that if $|M_S| > 2t$, then $\mathcal{I}$ is a No-instance, as the proof of the other case is symmetric to this one. Let us assume that $|M_S| > 2t$. Suppose, by way of contradiction, that $\mathcal{I}$ is a Yes-instance. Then, there exists a stable matching μ such that $\mathsf{balance}(\mu) \leq k$. Partition $M_S = M'_S \uplus M''_S$ as follows. Set M'_S to be the set of all m in M_S such that m is not the partner of $\mu(m)$ in μ_W.

$$M'_S = \{m \in M_S \mid f_{\mu(m)}(m) > f_{\mu(m)}(\mu_W(\mu(m)))\}.$$

Accordingly, set $M''_S = M_S \setminus M'_S$. Since $|M_S| > 2t$, at least one among $|M'_S|$ and $|M''_S|$ is (strictly) larger than t. Let us first handle the case where $|M'_S| > t$. Then,

$$\begin{aligned}\sum_{w\in W} f_w(\mu(w)) &= \sum_{\{w \mid \mu(w)\in M'_S\}} f_w(\mu(w)) + \sum_{\{w \mid \mu(w)\notin M'_S\}} f_w(\mu(w))\\ &\geq \sum_{\{w \mid \mu(w)\in M'_S\}} [f_w(\mu_W(w))+1] + \sum_{\{w \mid \mu(w)\notin M'_S\}} f_w(\mu_W(w))\\ &= \sum_{w\in W} f_w(\mu_W(w)) + \sum_{\{w \mid \mu(w)\in M'_S\}} 1 = O_W + |M'_S|\\ &> O_W + t \geq k.\end{aligned}$$

Here, the first inequality followed directly from the definition of M'_S. As we have reached a contradiction, it must hold that $|M''_S| > t$. However, we now have that

$$\begin{aligned}\sum_{m\in M} f_m(\mu(m)) &= \sum_{m\in M''_S} f_m(\mu(m)) + \sum_{m\notin M''_S} f_m(\mu(m))\\ &\geq \sum_{m\in M''_S} [f_m(\mu_M(m))+1] + \sum_{m\notin M''_S} f_m(\mu_M(m))\\ &= \sum_{m\in M} f_m(\mu_M(m)) + \sum_{\{m \mid m\in M''_S\}} 1 = O_M + |M''_S|\\ &> O_M + t \geq k.\end{aligned}$$

Here, the first inequality followed from the definition of M''_S. Indeed, for all $m \in M''_S$, we have that $f_{\mu(m)}(m) \leq f_{\mu(m)}(\mu_W(\mu(m)))$, else m would have belonged to M'_S. However, in this case we deduce that $m = \mu_W(\mu(m))$, and since $m \in M_S$, we have that $\mu(m) \neq \mu_M(m)$, which implies that $f_m(\mu(m)) \geq f_m(\mu_M(m)) + 1$. As we have again reached a contradiction, we conclude the proof. □

Marrying happy people. Towards the removal of happy people, we first need to handle the special case where there are no sad people. In this case, there is exactly one stable matching, which is the man-optimal stable matching (that is equal, in this case, to the woman-optimal stable matching). This immediately implies the safeness of the following rule.

Reduction Rule 4.5. *If $M_S = W_S = \emptyset$, then return* Yes *if* $\mathsf{balance}(\mu_M) \leq k$ *and* No *otherwise.*

Observation 4.2. *Reduction Rule 4.5 is safe.*

We now turn to discard happy people. When we perform this operation, we need to ensure that the balance of the instance is preserved. This is because we do not know which side (men or women) attains the $\mathsf{Bal}(\mathcal{I})$ value, hence we cannot reduce the quantity k by the dissatisfaction of the happy people on that side. Consequently, we need to ensure that $\mathsf{Bal}(\mathcal{I}) = \mathsf{Bal}(\mathcal{J})$, where $\mathcal{J}$ denotes the new instance resulting from the removal of some happy people. Towards this, we let (m_h, w_h) denote a *happy pair*, which is simply a pair of a happy man and happy woman who are matched to each other in every stable matching.[7] Then, we redefine the preference functions in a manner that allows us to transfer the "contributions" of m_h and w_h from $\mathsf{Bal}(\mathcal{I})$ to $\mathsf{Bal}(\mathcal{J})$ via some sad man and woman. We remark that these sad people exist because Reduction Rule 4.5 does not apply. The details are as follows.

Reduction Rule 4.6. *If there exists a happy pair (m_h, w_h), then proceed as follows. Select an arbitrary sad man m_s and an arbitrary sad woman w_s. Denote $M' = M \setminus \{m_h\}$ and $W' = W \setminus \{w_h\}$. For each person $a \in M' \cup W'$, the new preference function $f'_a : \mathcal{A}(a) \setminus \{m_h, w_h\} \to \mathbb{N}$ is defined as follows.*

- *For each $w \in \mathcal{A}(m_s) \setminus \{w_h\}$: $f'_{m_s}(w) = f_{m_s}(w) + f_{m_h}(w_h)$.*
- *For each $m \in \mathcal{A}(w_s) \setminus \{m_h\}$: $f'_{w_s}(m) = f_{w_s}(m) + f_{w_h}(m_h)$.*
- *For each $w \in W' \setminus \{w_s\}$: $f'_w = f_w|_{M'}$, and for each $m \in M' \setminus \{m_s\}$: $f'_m = f_m|_{W'}$.*

The new instance is $\mathcal{J} = (M', W', \{f'_m\}_{m\in M'}, \{f'_w\}_{w\in W'}, k)$.

The following lemma proves the forward direction of the safeness of Reduction Rule 4.6.

[7] Such a pair is often referred to as a fixed pair in the literature.

Lemma 4.8. *Let $\mu \in \mathrm{SM}(\mathcal{I})$ and $\mathcal{J}$ be the instance produced by Reduction Rule 4.6. Then, $\mu' = \mu \setminus \{(m_h, w_h)\}$ is a stable matching in $\mathcal{J}$ such that* $\mathsf{balance}_{\mathcal{J}}(\mu') = \mathsf{balance}_{\mathcal{I}}(\mu)$.

Proof. We first show that $\mu' \in \mathrm{SM}(\mathcal{J})$, i.e., μ' is stable in $\mathcal{J}$. Then, we show that $\mathsf{balance}_{\mathcal{J}}(\mu') = \mathsf{balance}_{\mathcal{I}}(\mu)$. By Reduction Rule 4.3, it holds that μ is a perfect matching in $\mathcal{I}$. Since (m_h, w_h) is a happy pair, it is clear that $(m_h, w_h) \in \mu$, and therefore μ' is a perfect matching in $\mathcal{J}$. Let $(m, w) \notin \mu'$ be some acceptable pair in $\mathcal{J}$. Since $\mu \in \mathrm{SM}(\mathcal{I})$ and it is a perfect matching, it holds that $f_m(w) > f_m(\mu(m))$ or $f_w(m) > f_w(\mu(w))$. Let us consider these two possibilities separately.

- Suppose that $f_m(w) > f_m(\mu(m))$. If $m \neq m_s$, then $f'_m(w) = f_m(w)$ and $f'_m(\mu'(m)) = f_m(\mu(m))$, and therefore $f'_m(w) > f'_m(\mu'(m))$. Else, $f'_m(w) = f_m(w) + f_{m_h}(w_h)$ and $f'_m(\mu'(m)) = f_m(\mu(m)) + f_{m_h}(w_h)$, and therefore again $f'_m(w) > f'_m(\mu'(m))$.
- Suppose that $f_w(m) > f_w(\mu(w))$. Analogously to the previous case, we get that $f'_w(m) > f'_w(\mu'(w))$.

Since the choice of (m, w) was arbitrary, we conclude that μ' does not have a blocking pair in $\mathcal{J}$, and therefore $\mu' \in \mathrm{SM}(\mathcal{J})$. To show that $\mathsf{balance}_{\mathcal{J}}(\mu') = \mathsf{balance}_{\mathcal{I}}(\mu)$, note that

$$\begin{aligned}
\mathsf{balance}_{\mathcal{J}}(\mu') &= \max\{\sum_{m \in M \setminus \{m_h\}} f'_m(\mu'(m)), \sum_{w \in W \setminus \{w_h\}} f'_w(\mu'(w))\} \\
&= \max\{f'_{m_s}(\mu'(m_s)) + \sum_{m \in M \setminus \{m_h, m_s\}} f'_m(\mu'(m)), \\
&\qquad f'_{w_s}(\mu'(w_s)) + \sum_{w \in W \setminus \{w_h, w_s\}} f'_w(\mu'(w))\} \\
&= \max\{f_{m_s}(\mu(m_s)) + f_{m_h}(w_h) + \sum_{m \in M \setminus \{m_h, m_s\}} f_m(\mu(m)), \\
&\qquad f_{w_s}(\mu(w_s)) + f_{w_h}(m_h) + \sum_{w \in W \setminus \{w_h, w_s\}} f_w(\mu(w))\} \\
&= \max\{\sum_{m \in M} f_m(\mu(m)), \sum_{w \in W} f_w(\mu(w))\} = \mathsf{balance}_{\mathcal{I}}(\mu).
\end{aligned}$$

This concludes the proof. □

The following observation helps to prove the reverse direction of the safeness of Reduction Rule 4.6.

Observation 4.3. *Let $\mathcal{I}$ be an instance of* Above-Min FBSM *on which Reduction Rules 4.1 to 4.2 have been exhaustively applied. Then, for every happy pair (m_h, w_h), it holds that $\mathcal{A}(m_h) = \{w_h\}$ and $\mathcal{A}(w_h) = \{m_h\}$.*

Lemma 4.9. *Let $\mu' \in \mathrm{SM}(\mathcal{J})$. Then, $\mu = \mu' \cup \{(m_h, w_h)\}$ is a stable matching in $\mathcal{I}$ such that* $\mathsf{balance}_{\mathcal{I}}(\mu) = \mathsf{balance}_{\mathcal{J}}(\mu')$.

Proof. We first show that $\mu \in \mathrm{SM}(\mathcal{I})$. By Reduction Rule 4.3, it holds that μ' is a perfect matching in $\mathcal{J}$. Since (m_h, w_h) is a happy pair and by Observation 4.3, we have that μ is a perfect matching in $\mathcal{I}$ such that neither m_h nor w_h participates in any pair that blocks μ (if such a pair exists). Let $(m, w) \notin \mu'$ be some acceptable pair in $\mathcal{I}$ such that $m \neq m_h$ and $w \neq w_h$. Since $\mu' \in \mathrm{SM}(\mathcal{J})$ and it is a perfect matching, it holds that $f'_m(w) > f'_m(\mu'(m))$ or $f'_w(m) > f'_w(\mu'(w))$. Let us consider these two possibilities separately.

- Suppose that $f'_m(w) > f'_m(\mu(m))$. If $m \neq m_s$, then $f_m(w) = f'_m(w)$ and $f_m(\mu(m)) = f'_m(\mu'(m))$, and therefore $f_m(w) > f_m(\mu(m))$. Else, $f_m(w) = f'_m(w) - f_{m_h}(w_h)$ and $f_m(\mu(m)) = f'_m(\mu'(m)) - f_{m_h}(w_h)$, and therefore again $f_m(w) > f_m(\mu(m))$.
- Suppose that $f'_w(m) > f'_w(\mu'(w))$. Analogously to the previous case, we get that $f_w(m) > f_w(\mu(w))$.

Since the choice of (m, w) was arbitrary, we conclude that μ does not have a blocking pair in $\mathcal{I}$, and therefore $\mu \in \mathrm{SM}(\mathcal{I})$. To show that $\mathsf{balance}_{\mathcal{I}}(\mu) = \mathsf{balance}_{\mathcal{J}}(\mu')$, we follow the exact same argument as the one present in the proof of Lemma 4.8. □

We now turn to justify the use of Reduction Rule 4.6. By Lemmata 4.8 and 4.9, it holds that $\mathsf{Bal}(\mathcal{I}) = \mathsf{Bal}(\mathcal{J})$, and that both $\mathsf{balance}_{\mathcal{I}}(\mu_M(\mathcal{I})) = \mathsf{balance}_{\mathcal{J}}(\mu_M(\mathcal{J}))$ and $\mathsf{balance}_{\mathcal{I}}(\mu_W(\mathcal{I})) = \mathsf{balance}_{\mathcal{J}}(\mu_W(\mathcal{J}))$. As the argument k remained untouched, we have the following lemma.

Lemma 4.10. *Reduction Rule 4.6 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Before we examine the preference functions more closely, we prove the following result.

Lemma 4.11. *Given an instance $\mathcal{I}$ of* Above-Min FBSM*, one can exhaustively apply Reduction Rules 4.1 to 4.6 in time $\mathcal{O}(\ell n^2)$ to obtain an instance $\mathcal{J}$ such that $t(\mathcal{J}) \leq t(\mathcal{I})$, where ℓ denotes the number of acceptable pairs and n denotes the number of agents. All people in $\mathcal{J}$ are sad and matched by every stable matching, and there exist at most $2t$ men and at most $2t$ women in $\mathcal{J}$.*

Proof. First, we analyze the time taken by each of the reduction rules. Notice that Reduction Rule 4.1 takes constant time. Using Proposition 2.1, in time $\mathcal{O}(\ell)$ we can compute the suffix and prefix for each agent. We assume that for each agent a, f_a is sorted. This can be performed in time $\mathcal{O}(n^2 \log n)$. In the beginning we spent time $\mathcal{O}(n^2 \log n + \ell) = \mathcal{O}(n^2 \log n)$ for the above two steps. Hence, in Reduction Rule 4.2, for an agent a, the $\mathsf{worst}(a)$ can be found in constant time. It takes time $\mathcal{O}(n)$ to update f_a. Hence, since Reduction Rule 4.2 can be applied on every agent, it takes time $\mathcal{O}(n^2)$. It can be easily verified that Reduction Rules 4.3 and 4.6 take time $\mathcal{O}(n^2)$. Whereas, Reduction Rules 4.4 and 4.5 can be executed in constant time. Hence, Reduction Rules 4.1 to 4.6 can be applied in time $\mathcal{O}(n^2)$. By applying them sequentially and exhaustively, the process either terminates by outputting a trivial Yes(No)-instance or it shrinks the size of the instance in each step. Therefore, it takes time $\mathcal{O}(\ell n^2)$ to apply the reduction rules exhaustively. Hence, the instance $\mathcal{J}$ is obtained in time $\mathcal{O}(n^2 \log n + \ell n^2) = \mathcal{O}(\ell n^2)$, and the claim that $t(\mathcal{J}) \leq t(\mathcal{I})$ follows from Lemmata 4.6 and 4.10. Due to Reduction Rules 4.3 and 4.6, we know that each person in $\mathcal{J}$ is sad and matched by every stable matching. Thus, due to Reduction Rule 4.4, we also know that there exist at most $2t$ men and at most $2t$ women in $\mathcal{J}$. □

Truncating high-values. So far we have bounded the number of people. However, the images of the preference functions can contain integers that are not bounded by a function polynomial in the parameter. Thus, even though the number of people is upper bounded by $4t$, the total size of the instance can be huge. Hence, we need to process the images of the preference functions.

The intuition behind Reduction Rule 4.7 (and thus the proof of Lemma 4.12) can be described as follows. Recall that we have already modified preference functions in Reduction Rule 4.2 so that they would not contain irrelevant information in the prefixes and suffixes. Our current goal is to truncate "high-values" of preference functions. Suppose that there exists a stable matching μ and a man m such that $f_m(\mu(m)) > t + f_m(\mu_M(m))$. That is, μ matches m to a woman whose functional value in f_m is larger than the functional value of the woman with whom m is matched by μ_M by at least t units. Then, $\mathsf{balance}(\mu) \geq \sum_{m' \in M} f_{m'}(\mu(m')) > O_M + t \geq k$. Hence, irrespective of whether or not the current instance is a Yes-instance, we know that μ is not a yes-certificate for our problem. We thus observe that we should delete all those acceptable pairs whose presence in any stable matching prevents its balance from being upper bounded by k. Formally, we have the following rule.

Reduction Rule 4.7. *If there exists an acceptable pair (m, w) such that $f_m(w) > (k - O_M) + f_m(\mu_M(m))$ or $f_w(m) > (k - O_W) + f_w(\mu_W(w))$, then define the preference functions as follows:*

- $f'_m = f_m|_{\mathcal{A}(m)\setminus\{w\}}$ *and* $f'_w = f_w|_{\mathcal{A}(w)\setminus\{m\}}$.
- *For all* $a \in M \cup W \setminus \{m, w\}$: $f'_a = f_a$.

The new instance is $\mathcal{J} = (M, W, \{f'_{m'}\}_{m' \in M}, \{f'_{w'}\}_{w' \in W}, k)$.

Lemma 4.12. *Reduction Rule 4.7 is safe, and $t(\mathcal{I}) \geq t(\mathcal{J})$.*

Proof. Without loss of generality, suppose that $f_m(w) > (k - O_M(\mathcal{I})) + f_m(\mu_M(m))$. Due to Reduction Rule 4.1, we have that $k - O_M(\mathcal{I}) \geq 0$, and therefore $f_m(w) > f_m(\mu_M(m))$, which implies that $w \neq \mu_M(m)$. That is, $(m, w) \notin \mu_M$. Since f_m and f'_m differ only at w, and f_w and f'_w differ only at m, we have that $\mu_M(\mathcal{I})$ is also a matching in $\mathcal{J}$, and due to Corollary 4.1, we deduce that $\mu_M(\mathcal{I}) = \mu_M(\mathcal{J})$. First, we would like to show that $t(\mathcal{I}) \geq t(\mathcal{J})$. For this purpose, it is sufficient to show that $O_M(\mathcal{I}) \leq O_M(\mathcal{J})$ and $O_W(\mathcal{I}) \leq O_W(\mathcal{J})$. Since $\mu_M(\mathcal{I}) = \mu_M(\mathcal{J})$, it is clear that $O_M(\mathcal{I}) = O_M(\mathcal{J})$. By Reduction Rule 4.3, $\mu_M(\mathcal{I})$ matches all people in $\mathcal{I}$. Thus, by Proposition 2.2 and since $\mu_M(\mathcal{I}) = \mu_M(\mathcal{J})$, we have that $\mu_W(\mathcal{J})$ matches all people in $\mathcal{J}$, and hence all people in $\mathcal{I}$. By Corollary 4.1, for any woman w' and the man m' most preferred by w' in $\mathcal{J}$, it holds that w' does not prefer m' over $\mu_W(w')$ in $\mathcal{I}$. Thus, by the definition of the preference functions, we have that $O_W(\mathcal{I}) \leq O_W(\mathcal{J})$.

To show that the rule is safe, we need to show that $\mathsf{Bal}(\mathcal{I}) \leq k$ if and only if $\mathsf{Bal}(\mathcal{J}) \leq k$. For this purpose, let us first suppose that $\mathsf{Bal}(\mathcal{I}) \leq k$. Then, there exists $\mu \in \mathsf{SM}(\mathcal{I})$ such that $\mathsf{balance}_{\mathcal{I}}(\mu) \leq k$. Notice that if $\mu(m) = w$, then since $f_m(w) > (k - O_M) + f_m(\mu_M(m))$ and by the equation (1) in "Perfect Matching", we have that

$$\begin{aligned} k &\geq \mathsf{balance}_{\mathcal{I}}(\mu) \geq \sum_{m' \in M} f_{m'}(\mu(m')) \\ &> (k - O_M(\mathcal{I})) + \sum_{m' \in M} f_{m'}(\mu_M(m')) = (k - O_M(\mathcal{I})) + O_M(\mathcal{I}), \end{aligned}$$

which is a contradiction. Hence, $(m, w) \notin \mu$. Therefore, μ is a matching in $\mathcal{J}$. By the definition of the new preference functions, if μ has a blocking pair in $\mathcal{J}$, then this pair also blocks μ in $\mathcal{I}$. Since μ is stable in $\mathcal{I}$, we have that μ is also stable in $\mathcal{J}$. Now, by our definition of the new preference functions, we have that $\mathsf{balance}_{\mathcal{I}}(\mu) = \mathsf{balance}_{\mathcal{J}}(\mu)$. Since $\mathsf{balance}_{\mathcal{I}}(\mu) \leq k$, we thus conclude that $\mathsf{Bal}(\mathcal{J}) \leq k$.

In the second direction, suppose that $\mathsf{Bal}(\mathcal{J}) \leq k$. Then, there exists $\mu \in \mathrm{SM}(\mathcal{J})$ such that $\mathsf{balance}_{\mathcal{J}}(\mu) \leq k$. Clearly, μ is also a matching in $\mathcal{I}$. Moreover, by the definition of the preference functions, every acceptable pair in $\mathcal{I}$ that is also present in $\mathcal{J}$ cannot block μ in $\mathcal{I}$, else it would have also blocked μ in $\mathcal{J}$. Thus, if μ has a blocking pair in $\mathcal{I}$, then this pair must be (m, w). We claim that (m, w) cannot block μ in $\mathcal{I}$, which would imply that $\mu \in \mathrm{SM}(\mathcal{I})$. Suppose, by way of contradiction, that this claim is not true. Recall that we have already proved that $\mu_M(\mathcal{I}) = \mu_M(\mathcal{J})$. Let us denote $\mu_M = \mu_M(\mathcal{I})$. We have that μ matches m, which implies that $f_m(\mu(m)) > f_m(w)$. Since $f_m(w) > (k - O_M(\mathcal{I})) + f_m(\mu_M(m))$, we deduce that $f_m(\mu(m)) > (k - O_M(\mathcal{I})) + f_m(\mu_M(m))$. Furthermore, since $f'_m(\mu(m)) = f_m(\mu(m))$, $f'_m(\mu_M(m)) = f_m(\mu_M(m))$ and $O_M(\mathcal{I}) = O_M(\mathcal{J})$, we get that $f'_m(\mu(m)) > (k - O_M(\mathcal{J})) + f'_m(\mu_M(m))$. However, we then have that

$$\begin{aligned} k &\geq \mathsf{balance}_{\mathcal{J}}(\mu) \geq \sum_{m' \in M} f'_{m'}(\mu(m')) \\ &> (k - O_M(\mathcal{J})) + \sum_{m' \in M} f'_{m'}(\mu_M(m')) = (k - O_M(\mathcal{J})) + O_M(\mathcal{J}), \end{aligned}$$

which is a contradiction. Therefore, $\mu \in \mathrm{SM}(\mathcal{I})$. The definition of the new preference functions imply that $\mathsf{balance}_{\mathcal{I}}(\mu) = \mathsf{balance}_{\mathcal{J}}(\mu)$. Since $\mathsf{balance}_{\mathcal{J}}(\mu) \leq k$, we may conclude that $\mathsf{Bal}(\mathcal{I}) \leq k$. □

Shrinking gaps. Currently, there might still exist a man $\overline{m}$ or a woman $\overline{w}$ such that $f_{\overline{m}}(\mu_M(\overline{m})) > 1$ or $f_{\overline{w}}(\mu_W(\overline{w})) > 1$, respectively. In the following rule, we would like to decrease some values assigned by the preference functions of such men and women in a manner that preserves equivalence.

Reduction Rule 4.8. *If there exist $\overline{m} \in M$ and $\overline{w} \in W$ such that $f_{\overline{m}}(\mu_M(\overline{m})) > 1$ and $f_{\overline{w}}(\mu_W(\overline{w})) > 1$, then define the preference functions as follows.*

- *For all $w \in \mathcal{A}(\overline{m})$: $f'_{\overline{m}}(w) = f_{\overline{m}}(w) - \alpha$ and for all $m \in \mathcal{A}(\overline{w})$: $f'_{\overline{w}}(m) = f_{\overline{w}}(m) - \alpha$, where $\alpha = \min\{f_{\overline{m}}(\mu_M(\overline{m})), f_{\overline{w}}(\mu_W(\overline{w}))\}$.*
- *For all $a \in M \cup W \setminus \{\overline{m}, \overline{w}\}$: $f'_a = f_a$.*

The new instance is $\mathcal{J} = (M, W, \{f'_{m'}\}_{m' \in M}, \{f'_{w'}\}_{w' \in W}, k - \alpha)$.

Lemma 4.13. *Reduction Rule 4.8 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Proof. Let us first observe that the set of acceptable pairs in $\mathcal{I}$ is equal to the set of acceptable pairs of $\mathcal{J}$. Furthermore, for every person a, including the cases where this person is either $\overline{m}$ or $\overline{w}$, any two acceptable partners b and b' of a that satisfy $f_a(b) < f_a(b')$ also satisfy $f'_a(b) < f'_a(b')$, and vice versa. Indeed, this observation follows directly from our definition of the new preference functions. In other words, if a person prefers some person over another in $\mathcal{I}$, then this person also has the same preference order in $\mathcal{J}$, and vice versa. We thus deduce that $\mathrm{SM}(\mathcal{I}) = \mathrm{SM}(\mathcal{J})$.

We proceed by claiming that for all $\mu \in \mathrm{SM}(\mathcal{I})$, we have that $\mathsf{balance}_{\mathcal{I}}(\mu) = \mathsf{balance}_{\mathcal{J}}(\mu) + \alpha$. Indeed, by the definition of the new preference functions and the equation (1) in "Perfect Matching", we have that

$$\begin{aligned} \mathsf{balance}_{\mathcal{I}}(\mu) &= \max\{\sum_{m \in M} f_m(\mu(m)), \sum_{w \in W} f_w(\mu(w))\} \\ &= \max\{f_{\overline{m}}(\mu(\overline{m})) + \sum_{m \in M \setminus \{\overline{m}\}} f_m(\mu(m)), f_{\overline{w}}(\mu(\overline{w})) + \sum_{w \in W \setminus \{\overline{w}\}} f_w(\mu(w))\} \\ &= \max\{f'_{\overline{m}}(\mu(\overline{m})) + \alpha + \sum_{m \in M \setminus \{\overline{m}\}} f'_m(\mu(m)), f'_{\overline{w}}(\mu(\overline{w})) + \alpha + \sum_{w \in W \setminus \{\overline{w}\}} f'_w(\mu(w))\} \\ &= \max\{\sum_{m \in M} f'_m(\mu(m)), \sum_{w \in W} f'_w(\mu(w))\} + \alpha = \mathsf{balance}_{\mathcal{J}}(\mu) + \alpha. \end{aligned}$$

Furthermore, the arguments above also show that $O_M(\mathcal{I}) = O_M(\mathcal{J}) + \alpha$ and $O_W(\mathcal{I}) = O_W(\mathcal{J}) + \alpha$. Hence, we have that $\mathsf{Bal}(\mathcal{I}) = \mathsf{Bal}(\mathcal{J}) + \alpha$. Since k was decreased by α, we may conclude that the rule is safe and that $t(\mathcal{I}) = t(\mathcal{J})$. □

After the exhaustive application of Reduction Rule 4.8, no more than $\mathcal{O}(t)$ number of times, at least one of the two parties does not have any member without a person assigned 1 by his/her preference function. As long as there exist a man and a woman whose preference lists do not have a person assigned 1, we can apply the reduction rule. Thus,

Observation 4.4. *Let $\mathcal{I}$ be an instance of* Above-Min FBSM *that is reduced with respect to Reduction Rules 4.1 to 4.8. Then, either* **(i)** *for every $m \in M$, we have that $f_m(\mu_M(m)) = 1$, or* **(ii)** *for every $w \in W$, we have that $f_w(\mu_W(w)) = 1$. In particular, either* **(i)** *$O_M = |M|$ or* **(ii)** *$O_W = |W|$.*

This concludes the description of our reduction rules. We are now ready to prove Lemma 4.1.

Proof of Lemma 4.1. Given an instance $\mathcal{I}$ of Above-Min FBSM, our kernelization algorithm exhaustively applies Reduction Rules 4.1 to 4.8, after which it outputs the resulting instance, $\mathcal{J}$, as the kernel. Notice that Reduction Rule 4.7 iterates over all (m, w) pairs to find one that satisfies the premise. This takes time $\mathcal{O}(n^2)$, and updating the preference functions takes time $\mathcal{O}(n)$, where n is the number of agents. Similar analysis shows that Reduction Rule 4.8 takes time $\mathcal{O}(n^2)$. Hence, using Lemma 4.11, each rule among Reduction Rules 4.1 to 4.8 can be applied in time $\mathcal{O}(n^2)$, and it either terminates the execution of the algorithm or shrinks the size of the instance. Hence, it is clear that the instance $\mathcal{J}$ is obtained in time $\mathcal{O}(\ell n^2)$, where ℓ is the number of acceptable pairs. The claims that $\mathcal{J}$ and $\mathcal{I}$ are equivalent and that $t(\mathcal{J}) \leq t(\mathcal{I})$ follow directly from the lemmata that prove the safeness of each rule as well as argue with respect to the parameter. By Lemma 4.11, we also have that the instance contains at most $2t$ (sad) men and at most $2t$ (sad) women.

It remains to show that the image of the preference function of each person is a subset of $\{1, 2, \ldots, t+1\}$. Since the domain of the preference function of each person is the set of acceptable partners for that person, it is sufficient to show that every acceptable pair (m, w) satisfies $f_m(w) \leq t+1$ and $f_w(m) \leq t+1$. By Reduction Rule 4.7, every acceptable pair (m, w) satisfies $f_m(w) \leq (k - O_M) + f_m(\mu_M(m))$ and $f_w(m) \leq (k - O_W) + f_w(\mu_W(w))$. Moreover, for any man m and woman w, it holds that $f_m(\mu_M(m)) \leq O_M - (|M| - 1)$ and $f_w(\mu_W(w)) \leq O_W - (|W| - 1)$. Thus, every acceptable pair (m, w) satisfies $f_m(w) \leq k - (|M| - 1)$ and $f_w(m) \leq k - (|W| - 1)$. By Reduction Rule 4.3 and Observation 4.4, we have that $t = k - \min\{O_M, O_W\} = k - |M| = k - |W|$. Hence, we further conclude that every acceptable pair (m, w) satisfies $f_m(w) \leq t+1$ and $f_w(m) \leq t+1$. Thus, the proof is complete. $\square$

4.2. Balanced stable marriage

Having proved Lemma 4.1, we have a kernel for Above-Min FBSM. We would like to employ this kernelization algorithm to design one for Above-Min BSM. For this purpose, we need to remove gaps from preference functions. Once we do this, we can view preference functions as preference lists and obtain the desired kernel. Hence, the following lemma concludes the proof of Theorem 2.

Lemma 4.14. Above-Min BSM *admits a kernel that has at most $3t$ men among whom at most $2t$ are sad and at most t are happy, at most $3t$ women among whom at most $2t$ are sad and at most t are happy. Additionally, each happy person has at most $2t + 1$ acceptable partners and each sad person has at most $t + 1$ acceptable partners. Moreover, every stable matching in the kernel is a perfect matching.*

In what follows, we describe our kernelization algorithm for Above-Min BSM. Let $\mathcal{K} = (M', W', \mathcal{L}_{M'}, \mathcal{L}_{W'}, k')$ be the input instance, which is an instance of Above-Min BSM. Our algorithm begins by applying the reduction given by Observation 2.1 to translate $\mathcal{K}$ into an instance $\mathcal{I}' = (M', W', \mathcal{F}_{M'}, \mathcal{F}_{W'}, k')$ of Above-Min FBSM. Then, our algorithm applies the kernelization algorithm given by Lemma 4.1 to $\mathcal{I}'$, obtaining a reduced instance $\mathcal{I} = (M, W, \mathcal{F}_M, \mathcal{F}_W, k)$ of Above-Min FBSM. By Lemma 4.1, this instance has at most $2t$ men, at most $2t$ women, and the image of the preference function of each person is a subset of $\{1, 2, \ldots, t+1\}$. To eliminate "gaps" in the preference functions, the algorithm proceeds as described below. Note that we no longer apply any reduction rule from Section 4.1 (even if its condition is satisfied), as we currently give a *new* kernelization procedure rather than an extension of the previous one. Let us first formally define the notion of a gap.

Definition 4.4. *Let $a \in M \cup W$, and i be a positive integer outside the image of f_a. If there exists an integer $j > i$ that belongs to the image of f_a, then f_a is said to have a* gap *at i.*

Inserting dummies. We have ensured that the largest number in the image of any preference function is at most $t + 1$. As every person is sad, every person must have at least two acceptable partners. Hence, it follows that there are at most $t - 1 \leq t$ gaps. To handle the gaps of *all* people, we create a set of t dummy men and t dummy women. Our objective is to introduce these dummy people as acceptable partners for people who have gaps in their preference functions, such that the function values of the dummy people would fill the gaps. Moreover, currently there are no happy people in the kernel, but after insertion the dummy people will be the happy people of the instance and create at most t happy pairs; and so the following rule would be applied only once.

Reduction Rule 4.9. *If there do not exist happy people, then let $X = \{x_1, x_2, \ldots, x_t\}$ denote a set of new (dummy) men, and $Y = \{y_1, y_2, \ldots, y_t\}$ denote a set of new (dummy) women. For each $i \in \{1, 2, \ldots, t\}$, initialize $\mathcal{A}(x_i) = \{y_i\}$, $\mathcal{A}(y_i) = \{x_i\}$ and $f_{x_i}(y_i) = f_{y_i}(x_i) = 1$. The new instance is $\mathcal{J} = (M \cup X, W \cup Y, \{f_m\}_{m \in M \cup X}, \{f_w\}_{w \in W \cup Y}, k + t)$.*

We note that for all $i \in \{1, 2, \ldots, t\}$, it holds that (x_i, y_i) is a happy pair.

Lemma 4.15. *Reduction Rule 4.9 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Proof. For all $i \in \{1, 2, \ldots, t\}$, it holds that (x_i, y_i) is a happy pair, and therefore it is present in every stable matching in $\mathcal{J}$. By our definition of the new preference functions, it is clear that if μ is a stable matching in $\mathcal{I}$, then $\mu' = \mu \cup \{(x_1, y_1), \ldots, (x_t, y_t)\}$ is a stable matching in $\mathcal{J}$. Moreover, if μ' is a stable matching in $\mathcal{J}$, then $\mu = \mu' \setminus \{(x_1, y_1), \ldots, (x_t, y_t)\}$ is a stable matching in $\mathcal{I}$. Hence, since for all $i \in \{1, 2, \ldots, t\}$, it holds that $f_{x_i}(y_i) = f_{y_i}(x_i) = 1$, our definition of the new preference functions directly implies that $\mathsf{Bal}(\mathcal{I}) + t = \mathsf{Bal}(\mathcal{J})$, $O_M(\mathcal{I}) + t = O_M(\mathcal{J})$ and $O_W(\mathcal{I}) + t = O_W(\mathcal{J})$. Hence, $t(\mathcal{I}) = t(\mathcal{J})$, which concludes the proof. $\square$

Reduction Rule 4.10. [Male version] *If there exists $m \in M$ such that f_m has a gap at some j, then select some $y_i \in Y \setminus \mathcal{A}(m)$, and set $\mathcal{A}'(m) = \mathcal{A}(m) \cup \{y_i\}$ and $\mathcal{A}'(y_i) = \mathcal{A}(y_i) \cup \{m\}$. The preference functions are defined as follows.*

- **The preference function of** m**:** *$f'_m(y_i) = j$, and for all $a \in \mathcal{A}(m)$, $f'_m(a) = f_m(a)$.*
- **The preference function of** y_i**:** *$f'_{y_i}(m) = \max\limits_{m' \in \mathcal{A}(y_i)} (f_{y_i}(m') + 1)$, and for all $a \in \mathcal{A}(y_i)$, $f'_{y_i}(a) = f_{y_i}(a)$.*
- *For all $a \in (M \cup W) \setminus \{m, y_i\}$: $f'_a = f_a$.*

The new instance is $\mathcal{J} = (M, W, \{f'_{m'}\}_{m' \in M}, \{f'_{w'}\}_{w' \in W}, k)$.

Lemma 4.16. *Reduction Rule 4.10 is safe, and $t(\mathcal{I}) = t(\mathcal{J})$.*

Proof. The only modifications that are performed are the insertion of m into the set of acceptable partners of y_i as the least preferred person, and the insertion of y_i into the set of acceptable partners of m in a location that previously contained a gap. Let us first observe that since $f_{x_i}(y_i) = f_{y_i}(x_i) = 1$ and $f'_{x_i}(y_i) = f'_{y_i}(x_i) = 1$, it holds that (x_i, y_i) is a happy pair in both $\mathcal{I}$ and $\mathcal{J}$. Hence, it is clear that $\mathsf{SM}(\mathcal{I}) = \mathsf{SM}(\mathcal{J})$, $O_M(\mathcal{I}) = O_M(\mathcal{J})$, $O_W(\mathcal{I}) = O_W(\mathcal{J})$, and that the balance of any stable matching in $\mathcal{I}$ is equal to its balance in $\mathcal{J}$. We thus conclude that the rule is safe and that $t(\mathcal{I}) = t(\mathcal{J})$. $\square$

Analogously, we have a female version of Reduction Rule 4.10 where we fill a gap in the preference function of some woman $w \in W$. Recall that we have $4t$ agents in the instance before applying Reduction Rules 4.9 and 4.10. Notice that Reduction Rule 4.9 takes time $\mathcal{O}(t)$. Since there are at most $\mathcal{O}(t)$ men with at most $\mathcal{O}(t)$ functional values, Reduction Rule 4.10 takes time $\mathcal{O}(t^2)$ to find a man with a gap in his preference list and to update the preference functions. Reduction Rule 4.9 is applied once. Reduction Rule 4.10 can be applied at most $\mathcal{O}(t^2)$ times since there are at most $\mathcal{O}(t^2)$ gaps. Hence, we take time $\mathcal{O}(t^4)$ to apply both the reduction rules exhaustively, that is, we take time polynomial in the input size. We do not repeat our arguments again, and straightaway state the following result, which follows directly from the safeness of Reduction Rule 4.9 and the male and female version of Reduction Rule 4.10.

Lemma 4.17. Above-Min FBSM *admits a kernel that has at most $3t$ men among whom at most $2t$ are sad, at most $3t$ women among whom at most $2t$ are sad, and such that each happy person has at most $2t + 1$ acceptable partners and each sad person has at most $t + 1$ acceptable partners. Moreover, the kernel contains at most t happy pairs and none of the preference functions contain any gaps.*

Finally, we translate the kernel for Above-Min FBSM to an instance of Above-Min BSM as follows. For all $a \in M \cup W$ and $b \in \mathcal{A}(a)$, we set $p_a(b) = f_a(b)$. The new instance is $\mathcal{J} = (M, W, \{p_m\}_{m \in M}, \{p_w\}_{w \in W}, k)$. Clearly, we thus obtain an equivalent instance in time $\mathcal{O}(t^2)$, which leads us to say that Lemma 4.14 is proved.

5. Parameterized algorithm

In this section, we design a parameterized algorithm for Above-Min BSM, and prove Theorem 3. As our algorithm is based on the method of bounded search trees, first we will give a brief description of this technique and then describe the procedure.

5.1. Bounded search tree: an overview

The running time of an algorithm that uses bounded search trees can be analyzed as follows (see, e.g., [10]). Suppose that the algorithm executes a branching rule which has ℓ branching options (each leading to a recursive call with the

Data: $\mathcal{I}, k$
Result: A matching μ such that $\mathsf{balance}(\mu) \le k$

for $M' \subseteq M_S$ **do**
  Set $r = k - O_M$.
  $\mathcal{F} = \texttt{Branch}(0, \emptyset)$.
  for $F \in \mathcal{F}$ **do**
    Let $\mu = F \cup \{(x_1, y_1), (x_2, y_2), \ldots, (x_h, y_h)\} \cup \{(m, \mu_M(m)) : m \in M_S \setminus M'\}$
    if *μ is a stable matching in $\mathcal{I}$* **then**
      if $\mathsf{balance}(\mu) \le k$ **then**
        return μ
      end
    end
  end
end
return $\emptyset$.

Algorithm 1: FPT Algorithm.

corresponding parameter value), such that in the i^{th} branch option the current value of the parameter decreases by b_i. Then, $(b_1, b_2, \ldots, b_\ell)$ is called the *branching vector* of this rule. For this branching vector the upper bound $T(k)$ on the number of leaves in the search tree is given by the following linear recurrence: $T(k) = T(k - b_1) + T(k - b_2) + \ldots + T(k - b_\ell)$, where k is the parameter.

We say that α is the *root* of $(b_1, b_2, \ldots, b_\ell)$ if it is the (unique) positive real root of $x^{b^*} = x^{b^* - b_1} + x^{b^* - b_2} + \cdots + x^{b^* - b_\ell}$, where $b^* = \max\{b_1, b_2, \ldots, b_\ell\}$. If $r > 0$ is the initial value of the parameter, and the algorithm (a) returns a result when (or before) the parameter is negative, and (b) only executes branching rules whose roots are bounded by a constant $c > 0$, then its running time is upper bounded by $\mathcal{O}^*(c^r)$. In particular, this yields a $\mathcal{O}^*(\alpha^r)$-time algorithm, where α is the root of the branching vector of our algorithm.

5.2. Description of the algorithm

Given an instance $\widehat{\mathcal{I}} = (\widehat{M}, \widehat{W}, \widehat{\mathcal{L}}_M, \widehat{\mathcal{L}}_W, \widehat{k})$ of Above-Min BSM, we begin by using the procedure given by Lemma 4.14 to obtain (in polynomial time) a kernel $\mathcal{I} = (M, W, \mathcal{L}_M, \mathcal{L}_W, k)$ of Above-Min BSM such that $\mathcal{I}$ has at most $3t$ men among whom at most $2t$ are sad, at most $3t$ women among whom at most $2t$ are sad. We denote the happy pairs in $\mathcal{I}$ by $(x_1, y_1), \ldots, (x_h, y_h)$ for some $h \le 3t$, where $X = \{x_1, \ldots, x_h\}$ and $Y = \{y_1, \ldots, y_h\}$. We denote the set of sad men by M_S, and for that we have $|M_S| \le 2t$.

We proceed by executing a loop where each iteration corresponds to a different subset $M' \subseteq M_S$. For a specific iteration, our goal is to determine whether there exists a stable matching μ such that the following conditions are satisfied: $\mathsf{balance}(\mu) \le k$; for all sad men $m \in M'$, $\mu(m) \ne \mu_M(m)$; and for all sad men $m \in M_S \setminus M'$, $\mu(m) = \mu_M(m)$. (Also, we recall that for any happy man x_i, we have $\mu(x_i) = \mu_M(x_i) = y_i$.) A stable matching satisfying these conditions (in the context of the current iteration) is said to be **valid**. We denote $r = k - O_M$, and observe that $r \le t$ because $t = k - \min\{O_M, O_W\}$.

Let us now consider some specific iteration. To determine whether there exists a valid stable matching, our plan is to execute a branching procedure, called `Branch` (depicted in Algorithm 2), which outputs every set F of mutually disjoint pairs of a man and a woman, where the man is in M' and the following conditions are satisfied.

1. Every man m in M' participates in exactly one pair (m, w) of F, and for that unique pair, it holds that $w \in \mathcal{A}(m)$ and $p_m(w) > p_m(\mu_M(m))$.
2. $\displaystyle\sum_{\substack{m \in M' \\ (m,w) \in F}} (p_m(w) - p_m(\mu_M(m))) \le r.$

5.3. The procedure `Branch`

We now present the description of the procedure `Branch` in the context of some set $M' \subseteq M_S$. Let us denote $M' = \{m_1, m_2, \ldots, m_p\}$ for an appropriate choice of p.

Each call to our procedure is of the form $\texttt{Branch}(i, \mathcal{P}, r)$ where $i \in \{0, 1, \ldots, p\}$, $\mathcal{P}$ is a set of pairs of a man in $\{m_1, m_2, \ldots, m_i\}$ and a (sad) woman, and $r \le k - O_M$, such that the following conditions are satisfied.

(I) Each man m in $\{m_1, m_2, \ldots, m_i\}$ participates in exactly one pair (m, w) of $\mathcal{P}$, and for that unique pair, it holds that $w \in \mathcal{A}(m)$ and $p_m(w) > p_m(\mu_M(m))$. The pairs in $\mathcal{P}$ are mutually disjoint. We define $\mu_{\mathcal{P}}$ as the *function* whose domain is $M_i = \{m_1, m_2, \ldots, m_i\}$ and which assigns to each man m in its domain the unique woman w such that $(m, w) \in \mathcal{P}$.

(II) Define $\mathsf{balance}(\mathcal{P}) = \displaystyle\sum_{m \in \{m_1, \ldots, m_i\}} (p_m(\mu_{\mathcal{P}}(m)) - p_m(\mu_M(m)))$. Then, $\mathsf{balance}(\mathcal{P}) \le (k - O_M)$.

Data: A pair $(i, \mathcal{P}, r)$ satisfying conditions (I) and (II)
Result: A family of sets of man-woman pairs that satisfy conditions (III) and (IV)

if $r < 0$ **then**
  return $\emptyset$.
else if $i = p$ **then**
  return $\{\emptyset\}$.
else
  Let $\widetilde{W} = \{w \in \mathcal{A}(m_{i+1}) \setminus Y : p_{m_{i+1}}(w) > p_{m_{i+1}}(\mu_M(m_{i+1}))\}$
  $\widetilde{W} = \widetilde{W} \setminus (\{w : (m, w) \in \mathcal{P}\} \cup \{\mu_M(m) : m \in M_S \setminus M'\})$
  if $|\widetilde{W}| < r$ **then**
    $W^* = \widetilde{W}$.
  else
    Let W^* be the set of r women in $\widetilde{W}$ who are most preferred by m_{i+1}.
  end
  Let $W^* = \{w_1, w_2, \ldots, w_q\}$.
  // Note that $q \leq r$.
  for $j \in \{1, 2, \ldots, q\}$ **do**
    Let $r' = r - (p_{m_{i+1}}(w_j) - p_{m_{i+1}}(\mu_M(m_{i+1})))$
    // since the increase in balance($\mathcal{P}$) when (m_{i+1}, w_j) is added to $\mathcal{P}$ is $p_{m_{i+1}}(w_j) - p_{m_{i+1}}(\mu_M(m_{i+1}))$, we decrease r by the same.
    Let $\mathcal{F}_j$= Branch$(i+1, \mathcal{P} \cup \{(m_{i+1}, w_j)\}, r')$.
    // $\mathcal{F}_j$ is the family of sets of pairs that is returned by the recursive call.
    $\mathcal{F} = \mathcal{F} \bigcup \{F \cup \{(m_{i+1}, w_j)\} : F \in \mathcal{F}_j\}$.
  end
end
return $\mathcal{F}$.

Algorithm 2: Branch$(i, \mathcal{P}, r)$.

Note that for the case $i = 0$, we have $\mathcal{P} = \emptyset$ and $r = k - O_M$, and the algorithm calls the procedure Branch$(0, \emptyset, r)$.

The objective of a call to Branch$(i, \mathcal{P}, r)$ is to return a family of sets, $\mathcal{F}$, where each set $F \in \mathcal{F}$ is a set of pairs of a man in $\{m_{i+1}, m_{i+2}, \ldots, m_p\}$ and a woman such that the following conditions are satisfied.

(III) Every man m in $\{m_{i+1}, m_{i+2}, \ldots, m_p\}$ participates in exactly one pair (m, w) of F, and for that unique pair, it holds that $w \in \mathcal{A}(m)$ and $p_m(w) > p_m(\mu_M(m))$. The pairs in F are mutually disjoint. Also, the pairs in F are disjoint from the pairs in $\mathcal{P}$. We define μ_F as the *function* whose domain is $\{m_{i+1}, m_{i+2}, \ldots, m_p\}$ and which assigns to each man m in its domain the unique woman w such that $(m, w) \in F$.

(IV) $\displaystyle\sum_{m \in \{m_{i+1}, m_{i+2}, \ldots, m_p\}} (p_m(\mu_F(m)) - p_m(\mu_M(m))) \leq r$ and $r = (k - O_M) - \text{balance}(\mathcal{P})$.

Thus, each member of the family $\mathcal{F}$ returned by Branch$(i, \mathcal{P}, r)$ extends the matching $\mu_{\mathcal{P}}$ such that the resulting matching satisfies our stated goal, condition 2 (page 40).

Measure: We use r as the measure to analyze the Branch procedure. The measure is initially equal to $k - O_M$ when $\mathcal{P} = \emptyset$. Therefore, according to the method of bounded search trees (see Section 5.1), in order to derive the $\mathcal{O}^*(2^r)$ running time, it is sufficient to ensure that Branch **(a)** returns a result when (or before) the measure r is negative, and **(b)** only executes branching rules whose roots are bounded by 2.

When the measure r is negative, we simply return $\mathcal{F} = \emptyset$, as there does not exist a set F satisfying the conditions above. Otherwise, when $i = p$, we return $\mathcal{F} = \{\emptyset\}$. The time complexity analysis is presented in the proof of Claim 5.2.

We describe the Branch procedure in words for the sake of exposition.

Overview of Algorithm 2: Consider a call Branch$(i, \mathcal{P}, r)$ where $r \geq 0$ and $i < p$. We define $\widetilde{W}$ to be the subset of sad women who are neither part of any pair in $\mathcal{P}$ nor are they matched to any man in $M_S \setminus M'$ under the man-optimal stable matching. That is, $\widetilde{W} = \{w \in \mathcal{A}(m_{i+1}) \setminus Y : p_{m_{i+1}}(w) > p_{m_{i+1}}(\mu_M(m_{i+1}))\} \setminus \big(\{w : (m, w) \in \mathcal{P}\} \cup \{\mu_M(m) : m \in M_S \setminus M'\}\big)$.

We further refine $\widetilde{W}$ by letting W^* denote the set of r women in $\widetilde{W}$ who are most preferred by m_{i+1}. In case $|\widetilde{W}| < r$, we simply denote $W^* = \widetilde{W}$. Let us also denote $W^* = \{w_1, w_2, \ldots, w_q\}$ for the appropriate value of $q \leq r$. Then, our procedure executes q branches. At the j^{th} branch, Branch calls itself recursively with $(i+1, \mathcal{P} \cup \{(m_{i+1}, w_j)\}, r')$. Eventually, Branch returns $\bigcup_{j=1}^{q} \{\{(m_{i+1}, w_j)\} \cup F : F \in \mathcal{F}_j\}$ where for each $j \in \{1, 2 \ldots, q\}$, we set $\mathcal{F}_j$ to be the family of sets of pairs that was returned by the recursive call of the j^{th} branch.

Correctness of the Branch procedure. The correctness follows from the observation that we are conducting an exhaustive search. More precisely, if there exists a set F satisfying Conditions (III) and (IV), then it must include exactly one of the pairs in $\{(m_{i+1}, w_j) : j \in \{1, 2, \ldots, q\}\}$, hence the manner in which the family $\mathcal{F}$ and the measure r are updated follows straightforwardly. We prove it formally in the following claim.

Claim 5.1. *Let $\mathcal{F}$ denote the output of Branch$(i, \mathcal{P}, r)$. Suppose that the matching $\mu_{\mathcal{P}}$ satisfies Conditions (I) and (II). Then each set $F \in \mathcal{F}$ outputted by Branch$(i, \mathcal{P}, r)$ yields a matching μ_F that satisfies Conditions (III) and (IV).*

Proof. Note that due to the condition in Line 4 of the algorithm, the output of the call Branch($p, \cdot, \cdot$) is $\emptyset$. Hence, a pair (m_{i+1}, w_j) is added to $\mathcal{F}$ if $p \geq i+1$ and $r \geq 0$. Due to Lines 7 and 16, the pairs are mutually disjoint. Since $M' = \{m_1, m_2, \ldots, m_p\}$, we have that $\mathcal{F}$ satisfies Condition (I).

In each iteration we add one (man, woman) pair to $\mathcal{P}$ and reduce r by the increase in $\mathsf{balance}(\mathcal{P})$. Due to Line 2 of the algorithm, the process stops if r becomes negative. Hence, $\mathsf{balance}(\mathcal{P}) \leq r$, that is, Condition (II) is satisfied.

For each man $m \in \{m_1, \ldots, m_p\}$, a pair involving m is added to F once at Line 17. Since the women already present in a pair in $\mathcal{P}$ are discarded in line 7, the pairs added to F are mutually disjoint. Also, if (m, w) is added to F, then $w \in W^\star$ which ensures that Condition (III) holds. In a branch, if we add (m_{i+1}, w_j) to F, we have already reduced r by $p_{m_{i+1}}(w_j) - p_{m_{i+1}}(\mu_M(m_{i+1}))$. Hence, Condition (IV) is maintained as an invariant. □

Claim 5.2. *Procedure* Branch *has a running time of* $\mathcal{O}^*(2^r)$.

Proof. As noted before, initially, the measure $r = (k - O_M) - \mathsf{balance}(\mathcal{P})$ and $r > 0$ because $\mathcal{P} = \emptyset$. We observe that at the j^{th} branch, the measure changes from $(k - O_M) - \mathsf{balance}(\mathcal{P})$ to $(k - O_M) - \mathsf{balance}(\mathcal{P} \cup \{(m_{i+1}, w_j)\})$. By our definition of w_j, we have $p_{m_{i+1}}(w_j) - p_{m_{i+1}}(\mu_M(m_{i+1}))) \geq j$. Hence, at the worst case, the branching vector is $(1, 2, \ldots, r)$.

Since the polynomial $x^r - \sum_{i=1}^{r-1} x^i - 1 = 0$ attains the global minimum at $x = \frac{2r}{r+1}$ which approaches 2 as $r \to \infty$, we can conclude that 2 is the best upper bound for the branching vector. □

Thus, the correctness and the time complexity of the procedure Branch is complete. □

5.4. Algorithm

Here, we describe the last step of the algorithm and we will argue that by having the branching procedure Branch, we can conclude the proof of the correctness of the algorithm.

We examine each set F in the outputted family of sets. Then, we check whether the pairs in F, together with $(x_1, y_1), (x_2, y_2), \ldots, (x_h, y_h)$ and every pair in $\{(m, \mu_M(m)) : m \in M_S \setminus M'\}$ form a stable matching whose balance is at most k. If the answer is positive, then we terminate the execution and accept. At the end, if we did not accept in any iteration, we reject.

Correctness. To see why the decision made in the above step is correct, suppose that there exists a stable matching μ whose balance is at most k. In this case, due to our exhaustive search, there exists some iteration in which μ is also valid. That iteration is associated with some $M' \subseteq M_S$. Observe that the set of pairs $\{(m, \mu(m)) : m \in M'\}$ is one of the sets in the outputted family $\mathcal{F}$. Indeed, the satisfaction of Condition 1 (page 40) follows from the fact that μ is a stable matching satisfying Conditions (I) and (III) of validity. Condition 2 (page 40) is satisfied because of the fact that μ satisfies Conditions (II) and (IV) of validity.

Claim 5.3. *The time complexity of our algorithm to solve* Above-Min BSM *is* $\mathcal{O}^*(8^t)$.

Proof. Let us denote by T the running time of the procedure Branch. Then, the total running time of our algorithm is bounded by $\mathcal{O}^*(2^{|M_S|} \cdot T) = \mathcal{O}^*(4^t \cdot T)$.

Recall that $r = k - O_M$, and $t = k - \min\{O_M, O_W\}$, hence $r \leq t$. Hence, to derive the running time in Theorem 3, it is sufficient to ensure that $T = \mathcal{O}^*(2^r)$, as proved in Claim 5.2. Thus, the time complexity of our algorithm is $\mathcal{O}^*(4^t \cdot 2^r) = \mathcal{O}^*(8^t)$ since $r \leq t$. □

Thus, Theorem 3 is proved. □

6. Conclusion

In this paper we study an optimization variant of the famous stable matching problem in the realm of parameterized complexity. Balanced Stable Matching is a constrained stable matching that lies in between the two extremes i.e., the man-optimal and the woman-optimal stable matchings, being globally desirable and fair to both sides. Mainly we study the problem with respect to two above-guarantee parameters. We show dichotomous results with respect to these two parameters. To show that the problem is FPT with respect to the first parameter, we design a kernelization algorithm for the problem instead of directly designing an FPT algorithm. We believe this idea can be used to show fixed parameter tractability for other NP-hard problems in this area.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Trends in Computational Social Choice, AI Access, 2017.
[2] David J. Abraham, Péter Biró, David F. Manlove, "Almost stable" matchings in the roommates problem, in: Proceedings of International Workshop on Approximation and Online Algorithms (WOAO), 2006, pp. 1–14.
[3] Nadja Betzler, A Multivariate Complexity Analysis of Voting Problems, PhD thesis, Friedrich-Schiller-Universität Jena, 2010.
[4] Nadja Betzler, Robert Bredereck, Jiehua Chen, Rolf Niedermeier, Studies in computational aspects of voting - a parameterized complexity perspective, in: The Multivariate Algorithmic Revolution and Beyond - Essays Dedicated to Michael R. Fellows on the Occasion of His 60th Birthday, 2012, pp. 318–363.
[5] Péter Biró, David F. Manlove, Eric J. McDermid, "Almost stable" matchings in the roommates problem with bounded preference list, Theor. Comput. Sci. 432 (2012) 10–20.
[6] Bernhard Bliem, Robert Bredereck, Rolf Niedermeier, Complexity of efficient and envy-free resource allocation: few agents, resources, or utility levels, in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16, 2016, pp. 102–108.
[7] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, Ariel D. Procaccia, Handbook of Computational Social Choice, Cambridge University Press, London, 2016.
[8] Robert Bredereck, Jiehua Chen, Piotr Faliszewski, Jiong Guo, Rolf Niedermeier, Gerhard Woeginger, Parameterized algorithmics for computational social choice: nine research challenges, Tsinghua Sci. Technol. 19 (4) (2014) 358.
[9] Jiehua Chen, Danny Hermelin, Manuel Sorge, Harel Yedidsion, How hard is it to satisfy (almost) all roommates?, in: 45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic, 2018, 35.
[10] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, Saket Saurabh, Parameterized Algorithms, Springer, 2015.
[11] Marek Cygan, Marcin Pilipczuk, Michal Pilipczuk, Jakub Onufry Wojtaszczyk, On multiway cut parameterized above lower bounds, ACM Trans. Algorithms 5 (1) (2013) 3.
[12] Rodney G. Downey, Michael R. Fellows, Fixed-parameter tractability and completeness II: on completeness for W[1], Theor. Comput. Sci. 141 (1&2) (1995) 109–131.
[13] Rodney G. Downey, Michael R. Fellows, Fundamentals of Parameterized Complexity, Texts in Computer Science, Springer, 2013.
[14] Piotr Faliszewski, Rolf Niedermeier, Parameterization in computational social choice, in: Encyclopedia of Algorithms, 2016, pp. 1516–1520.
[15] Tomás Feder, Stable networks and product graphs, PhD thesis, Stanford University, 1990.
[16] David Gale, Lloyd S. Shapley, College admissions and the stability of marriage, Am. Math. Mon. 69 (1962) 9–15.
[17] David Gale, Marilda Sotomayor, Some remarks on the stable matching problem, Discrete Appl. Math. 11 (4) (1985) 223–232.
[18] Shivam Garg, Geevarghese Philip, Raising the bar for vertex cover: fixed-parameter tractability above a higher guarantee, in: Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2016, pp. 1152–1166.
[19] Sushmita Gupta, Sanjukta Roy, Saket Saurabh, Meirav Zehavi, Group activity selection on graphs: parameterized analysis, in: Proceedings of 10th International Symposium Algorithmic Game Theory (SAGT), 2017, pp. 106–118.
[20] Sushmita Gupta, Saket Saurabh, Meirav Zehavi, On treewidth and stable marriage, CoRR, arXiv:1707.05404 [abs], 2017.
[21] Dan Gusfield, Three fast algorithms for four problems in stable marriage, SIAM J. Comput. 16 (1) (1987).
[22] Dan Gusfield, Robert W. Irving, The Stable Marriage Problem - Structure and Algorithms, Foundations of Computing Series, MIT Press, 1989.
[23] Koki Hamada, Kazuo Iwama, Shuichi Miyazaki, The hospitals/residents problem with quota lower bounds, in: Proceedings of the 19th Annual European Symposium on Algorithms (ESA), 2011, pp. 180–191.
[24] Ayumi Igarashi, Robert Bredereck, Edith Elkind, On parameterized complexity of group activity selection problems on social networks, in: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS), 2017, pp. 1575–1577.
[25] Ayumi Igarashi, Dominik Peters, Edith Elkind, Group activity selection on social networks, in: Proceedings of the 31st Conference on Artificial Intelligence (AAAI), 2017, pp. 565–571.
[26] Takao Inoshita, Robert W. Irving, Kazuo Iwama, Shuichi Miyazaki, Takashi Nagase, Improving man-optimal stable matchings by minimum changes to preference lists, Algorithms 6 (2) (2013) 371–382.
[27] Robert W. Irving, Kazuo Iwama, David F. Manlove, Shuichi Miyazaki, Yasufumi Morita, Hard variants of stable marriage, Theor. Comput. Sci. 276 (1–2) (2002) 261–279.
[28] Robert W. Irving, Paul Leather, Dan Gusfield, An efficient algorithm for the "optimal" stable marriage, J. ACM 34 (3) (1987) 532–543.
[29] Robert W. Irving, David F. Manlove, Gregg O'Malley, Stable marriage with ties and bounded length preference lists, J. Discret. Algorithms 7 (2) (2009) 213–219.
[30] Akiko Kato, Complexity of the sex-equal stable marriage problem, Jpn. J. Ind. Appl. Math. 10 (1) (1993) 1.
[31] Donald E. Knuth, Stable Marriage and Its Relation to Other Combinatorial Problems: An Introduction to the Mathematical Analysis of Algorithms, CRM Proceedings & Lecture Notes, American Mathematical Society, 1997.
[32] Hooyeon Lee, Vassilevska Williams, Complexity of the stable invitations problem, in: Proceedings of the 31st Conference on Artificial Intelligence (AAAI), 2017, pp. 579–585.
[33] Hooyeon Lee, Vassilevska Williams, Parameterized complexity of group activity selection, in: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS), 2017, pp. 353–361.
[34] Daniel Lokshtanov, N.S. Narayanaswamy, Venkatesh Raman, M.S. Ramanujan, Saket Saurabh, Faster parameterized algorithms using linear programming, ACM Trans. Algorithms 11 (2) (2014) 15.
[35] David F. Manlove, Algorithmics of Matching Under Preferences, Series on Theoretical Computer Science, vol. 2, WorldScientific, 2013.
[36] Dániel Marx, Ildikó Schlotter, Parameterized complexity and local search approaches for the stable marriage problem with ties, Algorithmica 58 (1) (2010) 170–187.
[37] Dániel Marx, Ildikó Schlotter, Stable assignment with couples: parameterized complexity and local search, Discrete Optim. 8 (1) (2011) 25–40.
[38] Eric McDermid, in: Personal Communications Between McDermid and Manlove, 2010.
[39] Eric McDermid, Robert W. Irving, Sex-equal stable matchings: complexity and exact algorithms, Algorithmica 68 (3) (2014) 545–570.
[40] Kitty Meeks, Baharak Rastegari, Stable marriage with groups of similar agents, in: Web and Internet Economics - 14th International Conference, WINE 2018, Oxford, UK, December 15-17, 2018, Proceedings, 2018, pp. 312–326.
[41] Matthias Mnich, Ildikó Schlotter, Stable marriage with covering constraints-a complete computational trichotomy, in: Proceedings of the 10th International Symposium of Algorithmic Game Theory (SAGT), 2017, pp. 320–332.
[42] Greg O'Malley, Algorithmic aspects of stable matching problems, PhD thesis, University of Glasgow, 2007.
[43] Maria S. Pini, Francesca Rossi, Brent Venable, Toby Walsh, Stable marriage problems with quantitative preferences, arXiv preprint, arXiv:1007.5120, 2010.
[44] Venkatesh Raman, M.S. Ramanujan, Saket Saurabh, Paths, flowers and vertex cover, in: Proceedings of the 19th Annual European Symposium on Algorithms (ESA), 2011, pp. 382–393.
[45] Alvin E. Roth, On the allocation of residents to rural hospitals: a general property of two-sided matching markets, Econometrica 54 (2) (1986) 425–427.