



# A stable marriage requires communication <sup>☆</sup>

Yannai A. Gonczarowski <sup>a,\*,1</sup>, Noam Nisan <sup>b,2</sup>, Rafail Ostrovsky <sup>c</sup>,  
Will Rosenbaum <sup>d,3</sup>

<sup>a</sup> Tel Aviv University, Blavatnik School of Computer Science, Israel

<sup>b</sup> The Hebrew University of Jerusalem, Rachel & Selim Benin School of Computer Science & Engineering and Federmann Center for the Study of Rationality, Israel

<sup>c</sup> University of California Los Angeles, Department of Computer Science and Mathematics, United States of America

<sup>d</sup> Max Planck Institute for Informatics, Germany

## ARTICLE INFO

### Article history:

Received 1 September 2016

Available online 29 July 2019

### Keywords:

Stable marriage

Stable matching

Approximately stable

Communication complexity

Distance to stability

## ABSTRACT

In 1976, Knuth asked whether the worst-case running-time of the Gale-Shapley algorithm for the Stable Marriage Problem can be improved when non-sequential access to the input is allowed. Partial negative answers were given by Ng and Hirschberg and as part of Segal's general communication-complexity analysis. We give a far simpler, yet significantly more powerful, argument showing that  $\Omega(n^2)$  Boolean queries of any type are required for finding a stable – or even approximately stable – marriage. Unlike Segal's lower bound, our lower bound generalizes additionally to (A) randomized algorithms, (B) allowing arbitrary separate preprocessing of the women's and men's respective preferences profiles, (C) related problems, e.g. whether a given pair is married in every/some stable marriage, (D) whether a proposed marriage is stable or far from stable. To analyze “approximately stable” marriages, we introduce the notion of “distance to stability” and provide an efficient algorithm for its computation.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

In the classic Stable Marriage Problem (Gale and Shapley, 1962), there are  $n$  **women** and  $n$  **men**; each woman has a full preference order over the men and each man has a full preference order over the women. The challenge is to find a **stable marriage**: a one-to-one mapping between women and men that is stable in the sense that it contains no **blocking pair**: a woman and man who mutually prefer each other over their current spouse in the marriage. Gale and Shapley (1962)

<sup>☆</sup> An extended abstract of this work appeared in the proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2015) (Gonczarowski et al., 2015).

\* Corresponding author.

E-mail addresses: [yannai@gonch.name](mailto:yannai@gonch.name) (Y.A. Gonczarowski), [noam.nisan@gmail.com](mailto:noam.nisan@gmail.com) (N. Nisan), [rafail@cs.ucla.edu](mailto:rafail@cs.ucla.edu) (R. Ostrovsky), [will.rosenbaum@gmail.com](mailto:will.rosenbaum@gmail.com) (W. Rosenbaum).

<sup>1</sup> This research was conducted while Gonczarowski was at the Hebrew University of Jerusalem (Einstein Institute of Mathematics, Rachel & Selim Benin School of Computer Science & Engineering and Federmann Center for the Study of Rationality) and Microsoft Research.

<sup>2</sup> This research was conducted in part while Nisan was at Microsoft Research.

<sup>3</sup> This research was conducted while Rosenbaum was at University of California Los Angeles (Department of Mathematics) and at Tel Aviv University (School of Electrical Engineering).

proved that such a stable marriage exists by providing an algorithm for finding one. Their algorithm takes  $\Theta(n^2)$  steps<sup>4</sup> in the worst case (Gale and Shapley, 1962), but only  $\Theta(n \log n)$  steps in the average case, over independently and uniformly chosen preferences (Wilson, 1972).

In 1976, Knuth asked (Knuth, 1976) whether this quadratic worst-case running-time can be improved upon. A related question was put forward in 1987 by Gusfield, who asked (Gusfield, 1987) whether even *verifying* the stability of a proposed marriage can be done any faster. As the input size here is quadratic in  $n$ , these questions only make sense in models that do not require sequentially reading the whole input, but rather provide some kind of random access to the preferences of the participants.

While Knuth's and Gusfield's questions arose from computational concerns, they also have tangible economic significance. In many real-world matching mechanisms, it is unreasonable to expect participants to provide their full preference list over all alternatives. This is not merely due to the effort of writing down an immense ordered list of alternatives, but also due to the sheer cognitive or physical effort of forming these preferences (for example, by conducting interviews). Formally, the process of forming or revealing one's preferences can be modeled as “querying” the individual's (perhaps implicitly defined) preferences. Each query consists of answering a single question about the individual's preferences that requires only a short response.<sup>5</sup> Thus, the inherent complexity of a marriage mechanism can be measured by the number of queries necessary to participate in the mechanism.

A partial answer to both Knuth's and Gusfield's questions was given by Ng and Hirschberg (1990), who considered a model that allows two types of unit-cost queries to the preferences of the participants: “what is woman  $w$ 's ranking of man  $m$ ?” (and, dually, “what is man  $m$ 's ranking of woman  $w$ ?”) and “which man does woman  $w$  rank at place  $k$ ?” (and, dually, “which woman does man  $m$  rank at place  $k$ ?”). In this model, they prove a tight  $\Theta(n^2)$  lower bound on the number of queries that any deterministic algorithm that solves the stable marriage problem, or even verifies whether a given marriage is stable, must make in the worst case. Chou and Lu (2010) later showed that even if one is allowed to separately query each of the  $\log n$  bits of the answer to queries such as “which man does woman  $w$  rank at place  $k$ ?” (and its dual query),  $\Theta(n^2 \log n)$  such Boolean queries are still required in order to deterministically find a stable marriage.

These results still leave two questions open. The first is whether some more powerful model may allow for faster algorithms. While many “natural” algorithms for stable marriage do fit into these models, there may be others that do not. Indeed, there exist problems for which “computationally unnatural” operations, such as various types of hashing, arithmetic operations, or even “cognitively natural” operations such as processing through a neural network, do give algorithmic speedups. Further, it may be the case that the participants' preferences are only defined implicitly by their actual input. For example, a participant's type could be a point in some (possibly high dimensional) geometric space, such that they prefer to be married to partners whose types are geometrically close to their own (see, for example, Bogomolnaia and Laslier, 2007; Bhatnagar et al., 2008). In this case, a natural query may be of the form “what is your type's  $k$ th coordinate?” Thus, it is of interest to ask whether an algorithm that queries the actual (geometric) input can be significantly more efficient than one that only queries the implicitly defined preferences.

The second question concerns randomized algorithms: can they do better than deterministic ones? This question is especially fitting for the stable marriage problem as the *expected* running-time is known to be small when the preferences are chosen uniformly at random.<sup>6</sup> We give a negative answer to both hopes, as well as several other related problems, thereby showing that answering a wide variety of basic questions related to stable marriages requires a quadratic number of queries (that is, requires querying nearly the entire preference structure):

**Theorem 1.1** (Informal, see Corollaries 3.1 and 3.2). *Any randomized (or deterministic) algorithm that uses any type of Boolean queries to the women's and to the men's preferences to solve any of the following problems requires  $\Omega(n^2)$  queries in the worst case:*

- (a) *finding an (approximately) stable marriage,*<sup>7</sup>
- (b) *determining whether a given marriage is stable or far from stable,*
- (c) *determining whether a given pair of participants is contained in some/every stable marriage,*
- (d) *finding any  $\varepsilon n$  pairs of participants that appear in some/every stable marriage.*

*These lower bounds hold even if we allow arbitrary preprocessing of all the men's preferences and of all the women's preferences separately. The lower bound for Part (a) holds regardless of which (stable or approximately stable) marriage is produced by the algorithm.*

<sup>4</sup> For a brief introduction to computer-science notation such as  $\Theta(n^2)$ ,  $\Omega(n^2)$ , and  $\mathcal{O}(n^2)$ , see Appendix A.

<sup>5</sup> The requirement that responses are “short” rules out, for example, the possibility of a participant being asked their entire preferences with a single query. In our analysis, we consider Boolean queries, i.e. queries are “yes/no” questions. Other models (e.g. that of Ng and Hirschberg, 1990) consider queries with slightly longer ( $\mathcal{O}(\log n)$ -bit) responses. This difference in the response size of allowed queries can only affect the complexity by at most a factor of the length (number of bits) of the longest response.

<sup>6</sup> In particular, this would be the case if the expected running time could be made small for *any* distribution on preferences, rather than just the uniform one.

<sup>7</sup> Our notion of “approximately stable marriage” is that the marriage shares many married couples with some stable marriage; see Definition 2.5 for a formal definition and a discussion, and Section 6 for proof of tractability of this notion.

Our proof of Theorem 1.1 comes from a reduction to the well-known lower bounds for the disjointness problem (Kalyanasundaram and Schintger, 1992; Razborov, 1992) in Yao's model (Yao, 1979) of two-party communication complexity (see Kushilevitz and Nisan, 1997, for a survey). We consider a scenario in which Alice holds the preferences of the  $n$  women and Bob holds the preferences of the  $n$  men, and show that each of the problems from Theorem 1.1 requires the exchange of  $\Omega(n^2)$  bits of communication between Alice and Bob.

We note that Segal (2007) shows by a general argument that any deterministic or nondeterministic<sup>8</sup> communication protocol among all  $2n$  participants for finding a stable marriage requires  $\Omega(n^2)$  bits of communication. Our argument for Theorem 1.1(a), in addition to being significantly simpler, generalizes Segal's result to account for randomized algorithms,<sup>9</sup> and even when considering only two-party communication between Alice and Bob (essentially allowing arbitrary communication within the set of women and within the set of men without cost). Furthermore, our lower bound holds even for merely determining whether a given marriage is stable or far from stable (Theorem 1.1(b)), as well as for the additional related problems described in Theorem 1.1(c,d). These results immediately imply the same lower bounds for any type of Boolean queries in the original computation model, as Boolean queries can be simulated by a communication protocol.

As indicated above, Theorem 1.1(a), as well as the corresponding lower bound on the two-party communication complexity, holds not only for stable marriages but also for approximately stable marriages. In the context of communication complexity, Chou and Lu (2010) also study such a relaxation of the stable marriage problem in a restricted computational model in which communication is non-interactive (a sketching model). Chou and Lu show that any (deterministic, non-interactive,  $2n$ -party) protocol that finds a marriage where only a constant fraction of participants are involved in blocking pairs requires  $\Theta(n^2 \log n)$  bits of communication. Our results are not directly comparable to these, as the two notions of approximate stability are not comparable. Furthermore, we use a significantly more general computation model (randomized, interactive, two-party), but give a slightly weaker lower bound.

Our lower bound for verification complexity (given in Theorem 1.1(b)) is tight. Indeed there exists a simple deterministic algorithm for verifying the stability of a proposed marriage, which requires  $\mathcal{O}(n^2)$  queries even in the weak comparison model that allows only for queries of the form “does woman  $w$  prefer man  $m_1$  over man  $m_2$ ?” and, dually, “does man  $m$  prefer woman  $w_1$  over woman  $w_2$ ?”<sup>10</sup> We do not know whether the lower bound is tight also for finding a stable marriage (Theorem 1.1(a)). Gale and Shapley's algorithm uses  $\mathcal{O}(n^2)$  queries in the worst case, but  $\mathcal{O}(n^2)$  of these queries require answers of  $\log n$  bits each. Thus, the algorithm requires a total of  $\mathcal{O}(n^2 \log n)$  Boolean queries, or bits of communication. We do not know whether  $\mathcal{O}(n^2)$  Boolean queries suffice for any algorithm. While the gap between Gale and Shapley's algorithm and our lower bound is small, we believe that it is interesting, as the number of queries performed by the algorithm is exactly linear in the input encoding length. An even slightly sublinear algorithm would therefore be interesting.<sup>11</sup> We indeed do not have any  $o(n^2 \log n)$  algorithm, even randomized and even in the strong two-party communication model, nor do we have any improved  $\omega(n^2)$  lower bound, even for deterministic algorithms and even in the simple comparison model.<sup>12</sup>

**Open Problem 1.1.** Consider the Comparison model for stable marriage that only allows for queries of the form “does man  $m$  prefer woman  $w_1$  over woman  $w_2$ ?” and, dually, “does woman  $w$  prefer man  $m_1$  over man  $m_2$ ?”. How many such queries are required, in the worst case, to find a stable marriage?

## 2. Model and preliminaries

### 2.1. The stable marriage problem

#### 2.1.1. Full preference lists

For ease of presentation, we consider a simplified version of the model of Gale and Shapley (1962). Let  $W$  and  $M$  be disjoint finite sets, of **women** and **men**, respectively, such that  $|W| = |M|$ .

<sup>8</sup> We use the term “nondeterministic,” as is customary in computer science, to refer not to randomized (i.e. probabilistic) algorithms, but rather to algorithms that, roughly speaking, need only verify the correctness of the output (rather than search for it). The precise definition, which is out of scope for this paper, is not required in order to follow the main text of this paper. See Kushilevitz and Nisan (1997) for a comparison of different models of communication complexity.

<sup>9</sup> We remark that in general, there may be an exponential gap between deterministic, nondeterministic, and randomized communication complexity.

<sup>10</sup> By simple batching, this verification algorithm can be converted into one that uses only  $\mathcal{O}(\frac{n^2}{\log n})$  queries, each of which returns an answer of length  $\log n$  bits (with each query still regarding the preferences of only a single participant). This highlights the fact that the lower bounds of Ng and Hirschberg (1990) crucially depend on the exact type of queries allowed in their model.

<sup>11</sup> Note that, as shown in Appendix E, the *nondeterministic* communication complexity is  $\Theta(n^2)$ , so proving higher lower bounds for the deterministic or randomized case may be challenging.

<sup>12</sup> It is interesting to note that Bei et al. (2013) identify a similar gap for the stable marriage problem in a dramatically different computation model of trial and error.

**Definition 2.1** (Full preferences).

1. A **full preference list** over  $M$  is a total ordering of  $M$ .
2. A **profile of full preference lists** for  $W$  over  $M$  is a specification of a full preference list over  $M$  for each woman  $w \in W$ . We denote the set of all profiles of full preference lists for  $W$  over  $M$  by  $\mathcal{F}(W, M)$ .
3. Given a profile  $P_W$  of full preference lists for  $W$  over  $M$ , a woman  $w \in W$  is said to **prefer a man  $m \in M$  over a man  $m' \in M$** , denoted by  $m \succ_w m'$ , if  $m$  precedes  $m'$  on the preference list of  $w$ . We say that  $w \in W$  **weakly prefers  $m$  over  $m'$**  if either  $m \succ_w m'$  or  $m = m'$ .

We define full preference lists over  $W$  and profiles of full preference lists for  $M$  over  $W$  analogously.

**Definition 2.2** (Perfect marriage). A **perfect marriage** between  $W$  and  $M$  is a one-to-one mapping between  $W$  and  $M$ .

**Definition 2.3** (Marriage market). A **marriage market** (with full preference lists) is a quadruplet  $(W, M, P_W, P_M)$ , where  $W$  and  $M$  are disjoint,  $|W| = |M|$ ,  $P_W \in \mathcal{F}(W, M)$  and  $P_M \in \mathcal{F}(M, W)$ .

**Definition 2.4** (Stability). Let  $(W, M, P_W, P_M)$  be a marriage market and let  $\mu$  be a perfect marriage (between  $W$  and  $M$ ).

1. A pair  $(w, m) \in W \times M$  is said to be a **blocking pair** (in  $(W, M, P_W, P_M)$ ) with respect to  $\mu$ , if each of  $w$  and  $m$  prefer the other over their spouse in  $\mu$ .
2.  $\mu$  is said to be **stable** if no blocking pairs exist with respect to  $\mu$ . Otherwise,  $\mu$  is said to be **unstable**.

### 2.1.2. Arbitrary preference lists

While our main results are phrased in terms of full preference lists and perfect marriages, some additional and intermediate results in Section 4 and in the appendix deal with an extended model, which allows for preferences to specify “blacklists” (i.e. declare some potential spouses as unacceptable) and for marriages to specify that some participants remain single. (This model is nonetheless also a simplified version of that of Gale and Shapley, 1962.) A (not necessarily full) **preference list** over  $M$  is a totally-ordered subset of  $M$ . We once again interpret a preference list as a ranking, from best to worst, of acceptable spouses. We interpret participants absent from a preference list as declared unacceptable, even at the cost of remaining single. Analogously, a **profile of preference lists** for  $W$  over  $M$  is a specification of a preference list over  $M$  for each woman  $w \in W$ ; we denote the set of all profiles of preference lists for  $W$  over  $M$  by  $\mathcal{P}(W, M) \supset \mathcal{F}(W, M)$ . In this extended model, a woman  $w$  is said to **prefer a man  $m$  over a man  $m'$**  not only when  $m$  precedes  $m'$  on the preference list of  $w$ , but also when  $m$  is on the preference list of  $w$  while  $m'$  is not. Again, if we say that  $w$  **weakly prefers  $m$  over  $m'$**  if either  $w$  prefers  $m$  over  $m'$  or  $m = m'$ . (We once again define preference lists and profiles of preference list for  $M$  over  $W$  analogously.)

A (not necessarily perfect) **marriage** between  $W$  and  $M$  is a one-to-one mapping between a subset of  $W$  and a subset of  $M$ . Given a marriage  $\mu$ , we denote the set of married women (i.e. the subset of  $W$  over which  $\mu$  is defined) by  $W_\mu$ ; we analogously denote the set of married men by  $M_\mu$ . For a marriage  $\mu$  to be **stable** (with respect to  $P_W$  and  $P_M$ ), we require not only that no blocking pair exist with respect to it, but also that no participant  $p \in W \cup M$  be married to someone not on the preference list of  $p$ .

We note that this model of arbitrary (not necessarily full) preference lists generalizes the model of full preference lists described in Section 2.1.1. Indeed, if the preference list of each participant happens to contain all participants of the opposite gender, then the two notions of stability agree. In particular, any marriage that is stable with respect to such preference lists prescribes that no participant remains single.

### 2.1.3. Known results

We now survey a few known results regarding the stable marriage problem, which we utilize throughout this paper. For the duration of this section, let  $(W, M, P_W, P_M)$  be a marriage market, defined either according to the definitions of Section 2.1.1 or according those of Section 2.1.2.

**Theorem 2.1** (Gale and Shapley, 1962). A stable marriage between  $W$  and  $M$  always exists. Moreover, there exists an  $M$ -optimal stable marriage, i.e. a stable marriage where each man weakly prefers his spouse in this stable marriage over his spouse in any other stable marriage.

Gale and Shapley (1962) provide an efficient algorithm for finding the  $M$ -optimal stable marriage. Their algorithm runs in  $\Theta(n^2)$  steps in the worst case, performing a query of  $\Theta(\log n)$  bits in each step. Hence, the Gale-Shapley algorithm queries  $\Theta(n^2 \log n)$  bits in the worst case.

**Theorem 2.2** (McVitie and Wilson, 1971). The  $M$ -optimal stable marriage is also the  $W$ -pessimal stable marriage, i.e. every other stable marriage is weakly preferred over it by each woman.

**Corollary 2.1** (*W-pessimal & M-pessimal  $\Rightarrow$  unique*). If a stable marriage is both the *W-pessimal stable marriage* and the *M-pessimal stable marriage*, then it is the unique stable marriage.

**Theorem 2.3** (*Roth's Rural Hospitals Theorem; Roth, 1986*).  $W_\mu$  (resp.  $M_\mu$ ) is the same for every stable marriage  $\mu$ .

#### 2.1.4. Approximately-stable marriages

In this section, we describe a notion of an “approximately stable marriage.” For ease of presentation, we restrict ourselves to marriage markets with full preference lists (i.e. the model described in Section 2.1.1). We define an approximately stable marriage as a perfect marriage that shares many married pairs with some (exactly) stable (perfect) marriage. Our definition is a natural generalization of that of Ünver (2005) (who considers only marriage markets with unique stable marriages), but it appears to be novel in its exact formulation. Our notion of approximate stability has the theoretical advantage of being derived from a metric on the set of all perfect marriages between  $W$  and  $M$ .

**Definition 2.5.** For any pair of perfect marriages  $\mu, \mu'$  between  $W$  and  $M$  (where  $|W| = |M| = n$ ), we define the **divorce distance** between  $\mu$  and  $\mu'$  to be<sup>13</sup>

$$d(\mu, \mu') = n - |\mu \cap \mu'|.$$

Note that  $d$  measures the minimum number of divorces required to convert  $\mu$  to  $\mu'$  (and vice versa). By abuse of notation, we denote the **divorce distance to stability** of a perfect marriage  $\mu$  to be

$$d(\mu) = \min_{\mu' \in \mathcal{M}} d(\mu, \mu')$$

where  $\mathcal{M}$  is the set of all stable perfect marriages between  $W$  and  $M$ . Thus,  $d(\mu)$  is the minimum number of divorces required to convert  $\mu$  into a stable marriage.

We say that a marriage  $\mu$  is  **$(1 - \varepsilon)$ -stable** if  $d(\mu) \leq \varepsilon n$ . We say  $\mu$  is  **$\varepsilon$ -unstable** if  $d(\mu) > \varepsilon n$ .

**Example 2.1.**  $d(\mu) = 0$  if and only if  $\mu$  is stable. Therefore, for  $\varepsilon = 0$  the concepts of 1-stability and 0-instability coincide precisely with (exact) stability and instability, respectively. Letting  $\varepsilon$  grow,  $(1 - \varepsilon)$ -stability is a weaker requirement for larger values of  $\varepsilon$ , while  $\varepsilon$ -instability is a stricter requirement for larger values of  $\varepsilon$ .

When only a single stable marriage exists (in this case, as noted above, our definition of approximate stability coincides with that of Ünver, 2005), efficiently computing the divorce distance to stability of a given perfect marriage  $\mu$  is straightforward: first use the Gale-Shapley algorithm to compute the  $M$ -optimal stable marriage (which in this case is the unique stable marriage), and then calculate the divorce distance between this (unique) stable marriage and  $\mu$ . Unfortunately, this computation fails to generalize. Brute-force computation of  $d(\mu)$  by iterating over all stable marriages is infeasible for general preferences, because the set of all stable marriages  $\mathcal{M}$  can be exponentially large (Knuth, 1976; Irving and Leather, 1986; Gusfield and Irving, 1989). Nonetheless, by exploiting the combinatorial structure of  $\mathcal{M}$ , we show that  $d(\mu)$  can still be efficiently computed (albeit in slightly slower time  $\tilde{O}(n^4)$ ). We describe an algorithm to this effect in Section 6. We believe this algorithm to be of independent interest.

The concept of divorce distance is perhaps most valuable in developing our understanding of the qualitative behavior of *exactly* stable marriage mechanisms. Given that (as our results on exact stability show) finding an exactly stable marriage requires very high communication (or alternatively, a very large number of queries), one may consider dynamic mechanisms that refine their output over time until reaching a stable marriage. Such mechanisms would produce some initial (not necessarily stable) marriage after an initial stage of communication/queries, and then after additional communication/queries, adjust the marriage to form a stable marriage.<sup>14</sup> If the social cost of each divorce due to this adjustment is high, then we would want to minimize the number of divorces in the second stage of the mechanism. That is, we would seek a marriage with small divorce distance to stability in the first stage, and would seek to replace it with the closest stable marriage in the second stage. The analysis of Section 6 implies that if such a first stage can be constructed, then a corresponding second stage can be implemented in a computationally efficient manner (using quadratically many queries, of course). Our main result regarding approximately stable marriages implies that such a first stage cannot be implemented using significantly less communication or fewer queries than finding an exactly stable marriage. We further discuss implications and interpretations of this result in Section 7.

<sup>13</sup> Abusing notation, we identify a perfect marriage  $\mu$  with the set of married pairs  $\{(w_1, \mu(w_1)), (w_2, \mu(w_2)), \dots\}$ . Thus,  $\mu \cap \mu'$  is the set of pairs (couples) that are married in both  $\mu$  and  $\mu'$ .

<sup>14</sup> In fact, one may argue that real-world marriages are formed in a similar manner: since no person can meet (and form preferences regarding) all of their potential spouses, they meet a relatively small number of potential spouses and marry based on the information gathered until that point. However, they do not stop meeting new people at that point in time. Over time, they may divorce their spouses in favor of spouses that they find more suitable.



**Comparison with blocking-pairs approximate stability** There does not appear to be consensus in the literature on precisely how to quantify the (in)stability of a marriage with respect to the participants' preferences. A common notion of approximate stability is the requirement for a marriage to have relatively few blocking pairs; see, e.g. Eriksson and Häggström, 2008. For example, one can define the **blocking-pairs instability** of a marriage  $\mu$ , which we denote by  $bp(\mu)$ , to be the number of blocking pairs, divided by  $n \cdot (n - 1)$  (where this divisor is the total number of “potential blocking pairs” in a perfect marriage, i.e. woman-man pairs who are not married to each other). As we now observe, having small divorce distance to stability is a strictly more demanding condition than having small blocking-pairs instability — that is, divorce distance to stability is a *finer* measure of instability than blocking-pairs instability.

**Proposition 2.1** (Small divorce distance implies, but is not implied by, few blocking pairs). Let  $W = \{w_1, \dots, w_n\}$  and  $M = \{m_1, \dots, m_n\}$ .

- (a) For every  $P_W$  and  $P_M$ , if a marriage  $\mu$  is  $(1 - \varepsilon)$ -stable for some  $\varepsilon$ , then  $bp(\mu) \leq 2\varepsilon$ .
- (b) There exist full preference lists  $P_W$  and  $P_M$ , and a perfect marriage  $\mu$ , such that  $\mu$  induces only a single blocking pair (so  $bp(\mu) = \frac{1}{n(n-1)}$ ), yet  $d(\mu) = n$  (so  $\mu$  is not  $(1 - \varepsilon)$ -unstable for any  $\varepsilon > 0$ ).

**Proof.** For Part (a), suppose that  $\mu$  is  $(1 - \varepsilon)$ -stable for some  $\varepsilon > 0$ . Therefore,  $d(\mu) = \varepsilon n$ , and by the definition of  $d(\mu)$  there is some stable marriage  $\mu'$  such that  $\mu$  and  $\mu'$  share  $(1 - \varepsilon)n$  pairs. Let  $W'$  and  $M'$  denote the sets of women and men, respectively, whose partners are the same in  $\mu$  and  $\mu'$ , so that  $|W'| = |M'| = (1 - \varepsilon)n$ . Since  $\mu'$  is stable, there are no blocking pairs  $(w, m) \in W' \times M'$  with respect to  $\mu'$ , and therefore there are no such pairs with respect to  $\mu$ . Therefore, any blocking pair  $(m, w)$  with respect to  $\mu$  must have  $w \notin W'$  or  $m \notin M'$  (or both). Since there are only  $\varepsilon n$  choices of such  $w$  (resp. of such  $m$ ), the total number of blocking pairs induced by  $\mu$  is less than  $2\varepsilon n(n - 1)$ , and so  $bp(\mu) \leq 2\varepsilon$ , as required.

For Part (b), we define full preference lists as follows. Each woman  $w_i$  prefers the men in order:  $m_1 \succ_{w_i} m_2 \succ_{w_i} \dots \succ_{w_i} m_n$ . Similarly, each man  $m_j$  for  $j \geq 2$  prefers the women in order:  $w_1 \succ_{m_j} w_2 \succ_{m_j} \dots \succ_{m_j} w_n$ . Finally, the preference list of  $m_1$  is defined as follows:

$$w_1 \succ_{m_1} w_n \succ_{m_1} w_2 \succ_{m_1} w_3 \succ_{m_1} \dots \succ_{m_1} w_{n-1}.$$

A straightforward argument shows that the marriage  $\mu' = \{(w_i, m_i) \mid i \in [n]\}$  is the unique stable marriage for these preference lists.<sup>15</sup> Consider now the marriage  $\mu$  defined by

$$\mu = \{(w_1, m_2), (w_2, m_3), \dots, (w_{n-1}, m_n), (w_n, m_1)\}.$$

The pair  $(w_1, m_1)$  is the only blocking pair induced by  $\mu$  (thus  $bp(\mu) = \frac{1}{n(n-1)}$ ). However,  $\mu \cap \mu' = \emptyset$ , so  $d(\mu) = n$ .  $\square$

The fact that the divorce distance is a finer measure of instability than blocking-pairs instability makes it easier for us to prove lower bounds on the complexity of finding  $(1 - \varepsilon)$ -stable marriages. (Indeed, we do not know if similar lower bounds apply to finding a marriage  $\mu$  satisfying  $bp(\mu) \leq \varepsilon$ ; see the discussion in Section 7.) Which of these two measures of instability is more appropriate, though, depends on the given context. One justification for the use of blocking-pairs instability is that  $bp(\mu)$  is exactly the probability that a uniformly random pair  $(w, m) \in (W \times M) \setminus \mu$  is a blocking pair. Thus, if  $bp(\mu)$  is small, it is perhaps unlikely that any instability will be detected, say, by randomly probing potential blocking pairs.

In many real-world matching markets, partnerships are formed with incomplete knowledge about the market, and the participants continue to learn about the market after an initial matching is formed. In such situations, it is reasonable to expect that participants will eventually discover instabilities if any are present in the initial matching. Depending on how the market responds to instabilities, the number of blocking pairs and the divorce distance can each be interpreted as a natural measure of the social cost of the instability. If the participants are obligated to remain in the initial matching, then blocking-pairs instability can be viewed as quantifying the total “suffering” of participants due to their awareness of being part of a blocking pair. If, on the other hand, participants are allowed to act on the instability, but must pay some cost to break existing partnerships, then the divorce distance to stability quantifies the (minimum) cost for achieving stability.

## 2.2. Communication complexity

We work in Yao's model (Yao, 1979) of two-party communication complexity (see Kushilevitz and Nisan, 1997, for a survey). Consider a scenario where two agents, Alice and Bob, hold values  $x$  and  $y$ , respectively, and wish to collaborate in performing some computation that depends on both  $x$  and  $y$ . Such a computation typically requires the exchange of some data between Alice and Bob. The **communication cost** of a given protocol (i.e. distributed algorithm) for such a computation is the number of bits that Alice and Bob exchange under this protocol in the worst case (i.e. for the worst  $(x, y)$ ); the

<sup>15</sup> For example, one can prove by induction on  $i$  that  $(w_1, m_1), (w_2, m_2), \dots, (w_i, m_i)$  must be in any stable marriage for  $i = 1, 2, \dots, n$ .

**communication complexity** of the computation that Alice and Bob wish to perform is the lowest communication cost of any protocol for this computation. Generalizing, we also consider **randomized** communication complexity, defined analogously using randomized protocols that for every given fixed input, produce a correct output with probability at least  $\frac{2}{3}$ .<sup>16</sup>

Of particular interest to us is the disjointness function, DISJ. Let  $n \in \mathbb{N}$  and let Alice and Bob hold subsets  $A, B \subseteq [n]$ , respectively. The value of the disjointness function is 1 if  $A \cap B = \emptyset$ , and 0 otherwise. We can also consider DISJ as a Boolean function by identifying  $A$  and  $B$  with their respective characteristic vectors  $\bar{x} = (x_i)_{i=1}^n$  and  $\bar{y} = (y_i)_{i=1}^n$ , defined by  $x_i = 1 \iff i \in A$  and  $y_j = 1 \iff j \in B$ . Thus, we can express DISJ using the Boolean formula  $\text{DISJ}(\bar{x}, \bar{y}) = \neg \bigvee_{i=1}^n (x_i \wedge y_i)$ . All of our results heavily rely on the following result of Kalyanasundaram and Schintger (1992) (see also Razborov, 1992):

**Theorem 2.4** (Communication complexity of DISJ; Kalyanasundaram and Schintger, 1992; Razborov, 1992). *Let  $n \in \mathbb{N}$ . The randomized (and deterministic) communication complexity of calculating  $\text{DISJ}(\bar{x}, \bar{y})$ , where  $\bar{x} \in \{0, 1\}^n$  is held by Alice and  $\bar{y} \in \{0, 1\}^n$  is held by Bob, is  $\Theta(n)$ . Further, this lower bound holds even for unique disjointness, i.e. if we require that the inputs  $\bar{x}$  and  $\bar{y}$  are either disjoint or uniquely intersecting:  $|\bar{x} \cap \bar{y}| \leq 1$ .*

Our results regarding lower bounds on communication complexities all follow from defining suitable **embeddings** of DISJ into various problems regarding stable marriages, i.e. mapping  $\bar{x}$  and  $\bar{y}$  into suitable marriage markets (more specifically, mapping  $\bar{x}$  into  $P_W$  and  $\bar{y}$  into  $P_M$ ), such that finding a stable marriage (or solving any of the other problems from Theorem 1.1) reveals the value of DISJ. Some of our proofs (namely those presented in Section 5) indeed assume that the input to DISJ satisfies  $|\bar{x} \cap \bar{y}| \leq 1$ .

### 3. Main results

The survey of our main results begins with the result that directly implies our answer to Knuth's question (Knuth, 1976) regarding the necessity of quadratic complexity for finding a stable marriage.

**Theorem 3.1** (Communication complexity of finding an approximately-stable marriage). *Let  $(W = \{w_1, \dots, w_n\}, M = \{m_1, \dots, m_n\}, P_W, P_M)$  be a marriage market with full preference lists, where  $P_W$  is held by Alice and  $P_M$  is held by Bob, and let  $0 \leq \varepsilon < \frac{1}{2}$ . The randomized (and deterministic) communication complexity of finding a  $(1 - \varepsilon)$ -stable marriage in  $(W, M, P_W, P_M)$  is  $\Omega(n^2)$ .*

We note that in particular, setting  $\varepsilon = 0$  in Theorem 3.1 shows that finding an (exactly) stable marriage requires  $\Omega(n^2)$  communication. Recall, though, that Knuth's question is phrased not with respect to communication complexity, but with respect to computational complexity. To answer Knuth's question, we reinterpret Theorem 3.1 using a standard argument that shows that the query complexity of a problem (that is, the worst-case number of Boolean queries any algorithm that solves this problem must make) is at least its communication complexity.

**Corollary 3.1** (Query complexity of finding an approximately-stable marriage). *Let  $(W = \{w_1, \dots, w_n\}, M = \{m_1, \dots, m_n\}, P_W, P_M)$  be a marriage market with full preference lists, and let  $0 \leq \varepsilon < \frac{1}{2}$ . Any randomized (or deterministic) algorithm that uses any type of Boolean queries to the women's and (separately) to the men's preferences to find a  $(1 - \varepsilon)$ -stable marriage requires  $\Omega(n^2)$  queries in the worst case.*

To see how Corollary 3.1 follows from Theorem 3.1, suppose there is a randomized algorithm  $A$  that computes a  $(1 - \varepsilon)$ -stable marriage using  $B$  Boolean queries to the women and men. We will use  $A$  to construct a  $B$ -bit communication protocol for the approximate stable marriage problem. The protocol works as follows. Alice and Bob both simulate  $A$ . Whenever  $A$  queries the women's preferences, Alice sends the result of the query to Bob (since Alice knows the women's preferences). Symmetrically, whenever  $A$  queries the men's preferences, Bob sends the result of the query to Alice. This protocol uses  $B$  bits of communication. Thus, by Theorem 3.1, we must have  $B = \Omega(n^2)$ , as desired.

Having given an answer (up to a factor of  $\log n$ ; see Open Question 1.1 in the introduction) to Knuth's question on the complexity of finding a stable marriage, one may ask how robust this answer is to variations in the studied problem. We have already shown that this answer is robust in one sense: the same lower bound applies also to finding approximately stable marriages. This shows that the hardness is not a consequence of some delicate phenomenon driven by the exact stability of the desired outcome. (Since any stable marriage is also  $(1 - \varepsilon)$ -stable, then allowing also approximately-stable marriage introduces additional possible solutions, so could have conceivably eased the task.) Another aspect of the robustness of our lower bound relates to Gusfield's question (Gusfield, 1987) regarding verifying the stability of a marriage: is it any easier to verify the correctness of a single concretely-proposed solution to the stable marriage problem? Setting an even higher bar of robustness, one can ask the following variant of Gusfield's question: say that we are given a marriage that is not a "borderline stable" marriage. That is, this marriage is guaranteed to be either exactly stable or very far from

<sup>16</sup> The results of this paper hold verbatim even if the constant  $\frac{2}{3}$  is replaced with any other fixed probability  $p$  with  $\frac{1}{2} < p \leq 1$ .

stable. Would it still be as hard to differentiate between these two cases assuming all “borderline” cases have been ruled out? Generalizing Knuth’s question in a different direction, one may also wonder if answering highly localized queries about the stable marriages in the market is any easier than finding a stable marriage. For example, is it easier to determine if a given woman-man pair is married in some (or in every) stable marriage?<sup>17</sup> We complement Theorem 3.1 by showing that our quadratic lower bound applies to any and all of the above, and even more.

**Theorem 3.2** (*Communication complexity of stable-marriage related problems*). Let  $(W = \{w_1, \dots, w_n\}, M = \{m_1, \dots, m_n\}, P_W, P_M)$  be a marriage market with full preference lists, where  $P_W$  is held by Alice and  $P_M$  is held by Bob. The randomized (and deterministic) communication complexity of solving any of the following problems is  $\Omega(n^2)$ .

- (a) determining whether a given marriage  $\mu$  is stable or  $\varepsilon$ -unstable, for fixed  $\varepsilon$  with  $0 \leq \varepsilon < 1$ .
- (b) determining whether a given pair of participants is contained in some/every stable marriage.
- (c) finding any  $\varepsilon n$  pairs that appear in some/every stable marriage, for fixed  $\varepsilon$  with  $0 < \varepsilon \leq 1$ .

Similarly to above, we note that setting  $\varepsilon = 0$  in Theorem 3.2(a) shows that verifying that a given marriage is (exactly) stable requires  $\Omega(n^2)$  communication. We also note that unlike the case of Theorem 3.1, where there is still a gap of  $\log n$  that prevents us from completely understanding the communication complexity of finding a stable marriage (see Open Question 1.1 in the introduction), this  $\Omega(n^2)$  lower-bound on the communication complexity of verifying the stability of a given marriage is tight: exhaustively querying all pairs of participants to naïvely check for the existence of a blocking pair requires  $\Theta(n^2)$  bits of communication in the worst case.<sup>18</sup> For the other problems analyzed in Theorem 3.2, a similar gap of  $\log n$  remains: it is possible to show that  $\mathcal{O}(n^2 \log n)$  communication suffices, due to a deterministic algorithm given by Gusfield (1987) for enumerating all pairs that belong to at least one stable marriage in  $\mathcal{O}(n^2 \log n)$  Boolean queries, but the question of a tight bound for these problems remains open as well.

As with Theorem 3.1 yielding Corollary 3.1, the same standard argument gives a query-complexity analogue of Theorem 3.2:

**Corollary 3.2** (*Query complexity of stable-marriage related problems*). Let  $(W = \{w_1, \dots, w_n\}, M = \{m_1, \dots, m_n\}, P_W, P_M)$  be a marriage market with full preference lists. Any randomized (or deterministic) algorithm that uses any type of Boolean queries to the women’s and (separately) to the men’s preferences to solve any of the following problems requires  $\Omega(n^2)$  queries in the worst case:

- (a) determining whether a given marriage  $\mu$  is stable or  $\varepsilon$ -unstable, for fixed  $\varepsilon$  with  $0 \leq \varepsilon < 1$ .
- (b) determining whether a given pair of participants is contained in some/every stable marriage.
- (c) finding any  $\varepsilon n$  pairs that appear in some/every stable marriage, for fixed  $\varepsilon$  with  $0 < \varepsilon \leq 1$ .

The variety of the problems covered in Theorems 3.1 and 3.2 conveys the main message of this paper: Roughly speaking, answering practically any meaningful basic question regarding a stable marriage in a marriage market, requires a quadratic number of queries, i.e. nearly amounts to querying the entire preference structure.

The proofs of Theorems 3.1 and 3.2 are given in Section 5.2. As we show there, all of these results can be proven by carefully embedding disjointness into a specially crafted marriage market that is described in Section 5.1. Before describing this embedding in full detail, we first describe in Section 4 a far more straightforward embedding of disjointness into a simpler marriage market. This embedding is just powerful enough to prove that verifying the (exact) stability of a given marriage requires  $\Omega(n^2)$  communication. We then add slightly more detail to this construction in order to prove that finding an (exactly) stable marriage also requires  $\Omega(n^2)$  communication. This gradual introduction to our machinery allows us to first expose the core of the argument that drives all of our main results (the fundamental relation between disjointness and stability) with only minimal surrounding detail. The simpler argument further allows readers who are unfamiliar with embeddings and disjointness to familiarize themselves with more straightforward examples of such techniques before delving into the intricacies of the general embedding.<sup>19</sup>

#### 4. Warm up: lower bounds for exact stability

In this section, we give direct proofs of our lower bounds for the communication complexity of finding or verifying an exactly stable marriage. We prove these lower bounds by embedding suitably large instances of DISJ into the problems of

<sup>17</sup> See Appendix C for analogous results also for another natural localized query that arises in the context of incomplete preference lists: determining whether a given participant is single in some (equivalently, in every) stable marriage.

<sup>18</sup> While Theorem 3.2(a) is phrased so that the marriage  $\mu$  is known by both Alice and Bob before the protocol commences, this tight bound still holds even if only one of them knows  $\mu$ , as the straightforward way of encoding a marriage between  $W$  and  $M$  requires  $\mathcal{O}(n \log n)$  bits.

<sup>19</sup> Furthermore, the proof in Section 4 of the lower bound for finding a stable marriage includes a novel application of the Rural Hospitals Theorem (Theorem 2.3) to facilitate the embedding of a decision problem (disjointness with no assumption of unique intersection) into the studied search problem, which we believe may be of independent interest.



finding a stable marriage or verifying the stability of some marriage. Thus, by Theorem 2.4 we obtain the desired lower bounds on the corresponding communication complexities. We note that the construction given in this section does not assume the input to DISJ to be uniquely intersecting.

**Definition 4.1.** Let  $n \in \mathbb{N}$ . We denote the set of pairs of distinct elements of  $\{1, \dots, n\}$  by  $[n]^2 = \{(i, j) \in \{1, \dots, n\}^2 \mid i \neq j\}$ . We note that  $|[n]^2| = n \cdot (n - 1)$ .

For the duration of this section, let  $n \in \mathbb{N}$ , and let  $W = \{w_1, \dots, w_n\}$  and  $M = \{m_1, \dots, m_n\}$  be disjoint sets such that  $|W| = |M| = n$ . Let  $\mu_{\text{id}}$  be the perfect marriage in which  $w_i$  is married to  $m_i$  for every  $i$ . To lower-bound the communication complexity of verifying the stability of a given marriage, we embed disjointness into verification of stability.

**Lemma 4.1** (Disjointness  $\hookrightarrow$  verifying stability). *There exist functions  $P_W : \{0, 1\}^{[n]^2} \rightarrow \mathcal{F}(W, M)$  and  $P_M : \{0, 1\}^{[n]^2} \rightarrow \mathcal{F}(M, W)$  such that for every  $\bar{x} = (x_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$  and  $\bar{y} = (y_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$ , the following are equivalent.*

- $\mu_{\text{id}}$  is stable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$
- $\text{DISJ}(\bar{x}, \bar{y}) = 1$ .

**Proof.** To define  $P_W(\bar{x})$ , for every  $i$  we define the preference list of  $w_i$  to consist of all  $m_j$  such that  $x_j^i = 1$ , in arbitrary order (say, sorted by  $j$ ), followed by  $m_i$ , followed by all other men in arbitrary order. Similarly, to define  $P_M(\bar{y})$ , for every  $j$  we define the preference list of  $m_j$  to consist of all  $w_i$  such that  $y_j^i = 1$ , in arbitrary order (say, sorted by  $i$ ), followed by  $w_j$ , followed by all other women in arbitrary order.

$\mu_{\text{id}}$  is unstable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y}) \iff$  there exist  $(i, j) \in [n]^2$  such that  $m_j \succ_{w_i} m_i$  and  $w_i \succ_{m_j} w_j \iff$  there exist  $(i, j) \in [n]^2$  such that  $x_j^i = 1$  and  $y_j^i = 1 \iff \text{DISJ}(\bar{x}, \bar{y}) = 0$ .  $\square$

**Remark 4.1.** A similar argument may be used to embed verification of stability back in disjointness.

To lower-bound the communication complexity of finding a stable marriage, we embed disjointness into finding a stable marriage through the intermediate problem of finding a stable marriage with respect to arbitrary (i.e. not necessarily full) preference lists. This embedding of disjointness into finding a stable marriage requires some more care than the above embedding into verification of stability, since unlike in Lemma 4.1 where we have embedded a decision problem (i.e. a problem for which the answer is Boolean) into a decision problem, here we embed a decision problem into a search problem (in this case, a problem for which the answer is a marriage). While in Section 5 we will mitigate this obstacle by assuming that the instance of disjointness that we embed is uniquely intersecting, here we overcome this obstacle via a novel application of the Rural Hospitals Theorem (Theorem 2.3), which we believe may be of independent interest.

**Lemma 4.2** (Disjointness  $\hookrightarrow$  finding a stable marriage (arbitrary preferences)). *There exist functions  $P_W : \{0, 1\}^{[n]^2} \rightarrow \mathcal{P}(W, M)$  and  $P_M : \{0, 1\}^{[n]^2} \rightarrow \mathcal{P}(M, W)$  such that for every  $\bar{x} = (x_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$  and  $\bar{y} = (y_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$ , both of the following hold.*

- a. If  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ , then  $\mu_{\text{id}}$  is the unique stable marriage with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$ .
- b. If  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ , then  $\mu_{\text{id}}$  is unstable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$ .

**Proof.** To define  $P_W(\bar{x})$ , for every  $i$  we define the preference list of  $w_i$  to consist of all  $m_j$  such that  $x_j^i = 1$ , in arbitrary order (say, sorted by  $j$ ), followed by  $m_i$  (with all other men absent). Similarly, to define  $P_M(\bar{y})$ , for every  $j$  we define the preference list of  $m_j$  to consist of all  $w_i$  such that  $y_j^i = 1$ , in arbitrary order (say, sorted by  $i$ ), followed by  $w_j$  (with all other women absent).

We first show that  $\mu_{\text{id}}$  is stable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$  iff  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ . Indeed, since every participant is married by  $\mu_{\text{id}}$  to someone on their preference list, we have:

$\mu_{\text{id}}$  is unstable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y}) \iff$  there exist  $(i, j) \in [n]^2$  such that  $m_j \succ_{w_i} m_i$  and  $w_i \succ_{m_j} w_j \iff$  there exist  $(i, j) \in [n]^2$  such that  $x_j^i = 1$  and  $y_j^i = 1 \iff \text{DISJ}(\bar{x}, \bar{y}) = 0$ .

It remains to show that if  $\mu_{\text{id}}$  is stable with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$ , then it is the unique stable marriage with respect to these profiles of preference lists. For the remainder of the proof assume, therefore, that  $\mu_{\text{id}}$  is stable (with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$ ). Let  $\mu$  be a stable marriage (with respect to these profiles of preference lists). As  $\mu_{\text{id}}$  is stable and perfect, by Theorem 2.3, since  $\mu$  is stable, it is perfect as well. Therefore, each  $p \in W \cup M$  is married by  $\mu$  to someone on the preference list of  $p$ , and so  $p$  weakly prefers  $\mu$  over  $\mu_{\text{id}}$ , as in the latter  $p$  is married to the last person on the preference list of  $p$ . Thus,  $\mu_{\text{id}}$  is both the  $W$ -pessimal stable marriage and the  $M$ -pessimal stable one, and so, by Corollary 2.1,  $\mu_{\text{id}}$  is the unique stable marriage.  $\square$

Our lower bound on the communication complexity of finding a stable marriage with respect to full preference lists follows from Lemma 4.2 by showing that we can embed the problem of finding a stable marriage with respect to possibly-partial preference lists into finding a stable marriage with respect to full preference lists. See Appendix B for details.

The techniques used to prove Lemmas 4.1 and 4.2 can also be used to prove Theorem 3.2(b) — see Appendix C. Although Theorem 3.2(b) shows that determining the marital status of a fixed pair  $(w, m)$  requires  $\Omega(n^2)$  communication, we do not know how to prove a similar lower bound for finding *some* married couple (see Open Problem 7.3 in Section 7). In the next section, we however show a weaker related result, namely that finding any constant fraction of the couples married in a stable marriage requires  $\Omega(n^2)$  communication (Theorem 3.2(c)). This result stems from a more elaborate embedding, which also yields our main results regarding approximate stability (Theorem 3.1 and Theorem 3.2(a)).

## 5. General proof of main results

### 5.1. Embedding disjointness into preferences

Similarly to the proofs given in Section 4, the proofs of the remaining results from Section 3 follow from embedding suitably large instances of DISJ into various problems regarding (approximately) stable marriages. In order to prove these remaining results, we reconstruct the embeddings to have the property that small changes in the participants' preferences yield very large changes in the global structure of the stable marriages for these preferences. Informally, we construct the preferences so that resolving blocking pairs resulting from such small changes in participants' preferences creates large rejection chains that ultimately affect most married couples.

#### 5.1.1. Preference description

Let  $n \in \mathbb{N}$  and let  $W$  and  $M$  be disjoint such that  $|W| = |M| = n$ . We divide the participants into three sets: **high**, **mid** and **low**, which we denote  $W_h$ ,  $W_m$  and  $W_l$  respectively for the women and  $M_h$ ,  $M_m$  and  $M_l$  respectively for the men. These sets have sizes

$$\begin{aligned} |W_h| &= |M_h| = \frac{1}{2}\delta n \\ |W_m| &= |M_m| = \frac{1}{2}(1 - \delta)n \\ |W_l| &= |M_l| = \frac{1}{2}n \end{aligned}$$

where  $\delta$  is a parameter with  $0 < \delta \leq 1$ , to be chosen later. The low and mid participants preferences will be fixed, while we will use the preferences of the high participants to embed an instance of disjointness of size  $(\delta n)^2/4$ . We assume that the participants are

$$W = \{w_1, w_2, \dots, w_n\}, \quad M = \{m_1, m_2, \dots, m_n\},$$

where in both cases the first  $\delta n/2$  participants are high, the next  $(1 - \delta)n/2$  participants are mid and the remaining  $n/2$  participants are low. Since the low and mid participants' preferences are the same for all instances, we describe those first. As before, the participants' preferences are symmetric in the sense that the men's and women's preferences are constructed analogously.

**low participants** The low women's preferences over men are “in order”:  $m_1 > m_2 > \dots > m_n$  (and symmetrically for low men, whose preference over women are “in order”). In particular, each low participant prefers all high participants over all mid participants over all low participants.

**mid participants** The mid participants prefer low participants over high participants over mid participants. Within each group, the preferences are “in order.” Specifically, the mid women have preferences  $m_{n/2+1} > m_{n/2+2} > \dots > m_n > m_1 > m_2 > \dots > m_{n/2}$ , and symmetrically for the men.

**high participants** We use the preferences of each of the high participants to encode a bit vector of length  $\delta n/2$ . Together, the men and women's preferences thus encode an instance of DISJ of size  $(\delta n)^2/4$ . For each  $w_i \in W_h$ , we denote her bit vector  $x_1^i, \dots, x_{\delta n/2}^i$ ; the preference list of  $w_i$ , from most-preferred to least-preferred, is:

1. men  $m_j \in M_h$  such that  $x_j^i = 1$ ;
2. men  $m \in M_l$ ;
3. men  $m \in M_m$ ;
4. men  $m_j \in M_h$  such that  $x_j^i = 0$ .

Within each group, the preferences are once again “in order”, i.e. sorted by numeric index. The men's preferences are constructed analogously, with each man  $m_j$  encoding the bit vector  $y_1^j, \dots, y_{\delta n/2}^j$  and preferring first and foremost women  $w_i \in W_h$  such that  $y_j^i = 1$ .

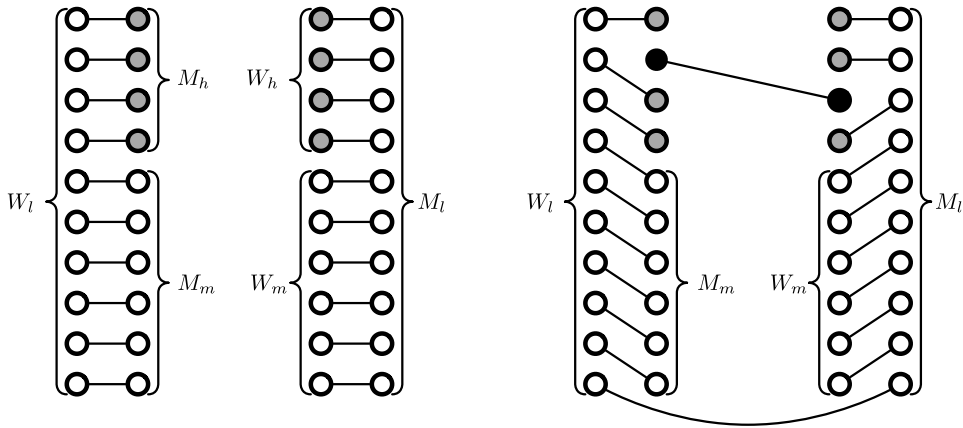


Fig. 1. The (unique) stable marriages  $\mu_1$  for disjoint (left) and  $\mu_0$  for uniquely-intersecting (right) instances of the preferences described in Section 5.1.1.

### 5.1.2. Stable marriage description

**Lemma 5.1.** Any instance of the stable marriage problem with preferences described above corresponding to  $\text{DISJ}(\bar{x}, \bar{y}) = 1$  has a unique stable marriage  $\mu_1$  given by (see the left side of Fig. 1)

$$\mu_1 = \{(m_i, w_{i+n/2}) \mid i = 1, 2, \dots, n/2\} \cup \{(m_{i+n/2}, w_i) \mid i = 1, 2, \dots, n/2\}.$$

**Proof.** Let  $\mu$  be a stable marriage; we will show that  $\mu = \mu_1$ . We first argue that every high and mid participant is married to a low participant in  $\mu$ . Suppose to the contrary that some  $w = w_i$  for  $i \leq n/2$  is married to some  $m = m_j$  with  $j \leq n/2$  in  $\mu$ . By the definition of the preferences and the assumption that  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ , at least one of  $w$  and  $m$  prefers every low participant over their spouse. Assume without loss of generality that  $w$  prefers all  $m' = m_{j'}$  with  $j' > n/2$  over  $m$ . That is,  $w$  prefers all low men over her spouse  $m$ . Since  $w$  is married to a medium or high man, there must be some low man  $m'$  that is married to a low woman  $w'$ . But  $m'$  prefers all high and medium women over  $w'$ . In particular, he prefers  $w$  over  $w'$ . Therefore,  $(w, m')$  is a blocking pair, so  $\mu$  is not stable. Thus any stable marriage must marry low participants to mid or high participants and vice versa.

Now we argue that if  $(w_i, m_{j+n/2}) \in \mu$ , then we must have  $i = j$ . The argument for pairs  $(w_{i+n/2}, m_j)$  is identical. Suppose that  $(w_i, m_{j+n/2}) \in \mu$  with  $i < j$ . Then there is some  $j' < j$  such that  $m' = m_{j'+n/2}$  is married to  $w' = w_{i'}$  with  $i' > i$ . But then  $(w_i, m')$  mutually prefer each other, contradicting the stability of  $\mu$ . We arrive at a similar contradiction if  $i > j$ , hence we must have  $i = j$ , as desired.  $\square$

**Lemma 5.2.** Suppose we have a stable marriage instance with preferences described above corresponding to  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ , with  $\bar{x}$  and  $\bar{y}$  uniquely intersecting. Let  $x_\beta^\alpha = y_\beta^\alpha = 1$  be the uniquely-intersecting entry of  $\bar{x}, \bar{y}$ . In this case, there exists a unique stable marriage  $\mu_0$  given by (see the right side of Fig. 1)

$$\begin{aligned} \mu_0 = & \{(w_\alpha, m_\beta)\} \cup \{(w_i, m_{i+n/2}) \mid i < \alpha\} \\ & \cup \{(w_{i+n/2}, m_i) \mid i < \beta\} \\ & \cup \{(w_i, m_{i+n/2-1}), \alpha < i \leq n/2\} \\ & \cup \{(w_{i+n/2-1}, m_i), \beta < i \leq n/2\} \cup \{(w_n, m_n)\}. \end{aligned}$$

**Proof.** We first argue that  $(w_\alpha, m_\beta) \in \mu$  for any stable marriage  $\mu$  for the preferences described above. Since  $\mu$  is stable, if  $(m_\alpha, w_\beta) \notin \mu$ , then at least one of  $w_\alpha$  and  $m_\beta$ , say  $w_\alpha$ , must be married to someone she prefers over  $m_\beta$ . From  $w_\alpha$ 's preferences, this implies that  $(w_\alpha, m) \in \mu$  for some  $m = m_j$  with  $j < \beta$  for which  $x_j^\alpha = 1$ . Since the instance of DISJ is uniquely intersecting, we must have  $y_j^\alpha = 0$ . Thus  $m$  prefers all low women over  $w_\alpha$ . Since at most  $n/2 - 1$  medium and high men are married to low women (indeed  $m$  is a high man married to a high woman) and there are  $n/2$  low women, some low woman  $w$  is married to a low man. But then  $w$  and  $m$  mutually prefer each other, hence forming a blocking pair. Thus, we must have  $(w_\alpha, m_\beta) \in \mu$ .

The remainder of the proof of the lemma is analogous to the proof of Lemma 5.1 if we remove  $w_\alpha$  and  $m_\beta$  from all the participants' preferences.  $\square$

**Lemma 5.3.** *The marriages  $\mu_0$  and  $\mu_1$  from the previous two lemmas satisfy  $d(\mu_0, \mu_1) \geq (1 - \delta)n$ .*

**Proof.** This follows from the following two observations:

1. All mid women and men  $M_m \cup W_m$  have different spouses in  $\mu_0$  and  $\mu_1$ .
2. No mid women are married to mid men in either  $\mu_0$  or  $\mu_1$ .

From these facts, we can conclude that  $d(\mu_0, \mu_1) = n - |\mu_0 \cap \mu_1| \geq |W_m| + |M_m| = (1 - \delta)n$ .  $\square$

## 5.2. Derivation of main results

In this section we use the construction of Section 5.1 to prove Theorems 3.1 and 3.2 from Section 3.

**Proof of Theorem 3.1.** Suppose that  $\Pi$  is a randomized communication protocol (between Alice and Bob) that outputs a  $(1 - \varepsilon)$ -stable marriage  $\mu$  using  $B$  bits of communication. As  $\varepsilon < 1/2$ , there exists  $\delta$  sufficiently small such that  $\varepsilon < (1 - \delta)/2$ . Suppose  $\Pi$  outputs a  $(1 - \varepsilon)$ -stable marriage  $\mu$  for the preferences described in Section 5.1.1. If  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ , then by Lemma 5.1,  $\mu_1$  is the unique stable marriage, so  $d(\mu, \mu_1) \leq \varepsilon n$ .

Suppose  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ . By Lemma 5.2,  $\mu_0$  is the unique stable marriage, so  $d(\mu, \mu_0) \leq \varepsilon n < (1 - \delta)n/2$ . Applying Lemma 5.3 and the triangle inequality, we obtain  $d(\mu_1, \mu) > (1 - \delta)n/2 > \varepsilon n$ . Thus, if  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ , then  $d(\mu, \mu_1) < \varepsilon n$  and if  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ , then  $d(\mu, \mu_1) > \varepsilon n$ . Given  $\mu$ , Alice and Bob can compute  $d(\mu, \mu_1)$  without communication, so they can use  $\Pi$  to determine the value of  $\text{DISJ}(\bar{x}, \bar{y})$  using  $B$  bits of communication. Thus,  $B = \Omega(n^2)$  by Theorem 2.4, as desired.  $\square$

**Proof of Theorem 3.2.** For Part (a), suppose that  $\Pi$  is a randomized communication protocol that determines whether a given marriage  $\mu$  is stable or  $\varepsilon$ -unstable with respect to given preferences using  $B$  bits of communication. As  $\varepsilon < 1$ , there exists  $\delta$  sufficiently small such that  $1 - \delta > \varepsilon$ . Let  $\mu_1$  be the marriage defined in Lemma 5.1; by that lemma, if  $\text{DISJ}(\bar{x}, \bar{y}) = 1$ , then  $\mu_1$  is stable (with respect to the preferences described in Section 5.1.1). By Lemmas 5.2 and 5.3, if  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ , then  $\mu_1$  is  $\varepsilon$ -unstable. Thus, if  $\Pi$  determines whether  $\mu_1$  is stable or  $\varepsilon$ -unstable, then  $\Pi$  also determines the value of  $\text{DISJ}(\bar{x}, \bar{y})$ , hence  $B = \Omega(n^2)$  by Theorem 2.4.

For Part (b), suppose that  $\Pi$  is a randomized communication protocol that for a given pair  $(w, m)$  determines whether  $(w, m) \in \mu$  for some (every) stable marriage  $\mu$  using  $B$  bits of communication. Set  $\delta = 1$ . By choosing preferences as in Section 5.1.1 and taking  $(w, m) = (w_n, m_n)$ , by Lemmas 5.1 and 5.2,  $(w, m)$  is in some (equivalently every) stable marriage for the given preferences if and only if  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ . Thus, once again by Theorem 2.4,  $B = \Omega(n^2)$ .

Finally, for Part (c), suppose that  $\Pi$  is a randomized communication protocol that outputs  $\varepsilon n$  pairs contained in some (every) stable marriage using  $B$  bits of communication. Choose preferences as described in the Section 5.1.1 with some  $0 < \delta < \varepsilon$ , say  $\delta = \varepsilon/2$ . Recall from the proof of Lemma 5.3 that no participants in  $W_m$  and  $M_m$  are ever married to one another in a stable marriage. Therefore, since  $|W_m| + |M_m| = (1 - \delta)n > (1 - \varepsilon)n$  and since  $\Pi$  outputs  $\varepsilon n$  pairs, we have that  $\Pi$  must output some pair  $(w, m)$  with  $w \in W_m$  or  $m \in M_m$ . Recall from the proof of Lemma 5.3 that knowing the stable spouse of any participant in  $W_m$  or  $M_m$  reveals the value of  $\text{DISJ}(\bar{x}, \bar{y})$ . Thus, by Theorem 2.4,  $B = \Omega(n^2)$ .  $\square$

## 6. Computing distance to stability

In this section, we describe an efficient method for computing the divorce distance to stability of a given marriage,  $\mu$ . Recall that the divorce distance to stability is given by

$$d(\mu) = \min_{\mu' \in \mathcal{M}} d(\mu, \mu') \quad \text{where} \quad d(\mu, \mu') = n - |\mu \cap \mu'|.$$

In fact, our algorithm solves the following general problem, which we believe may be of independent interest: given a marriage market  $(W, M, P_W, P_M)$  and an arbitrary marriage  $\mu$ , find a stable marriage  $\mu'$  that shares the greatest number of pairs with  $\mu$ . Since the set  $\mathcal{M}$  of all stable marriages can be exponentially large (Knuth, 1976; Irving and Leather, 1986; Gusfield and Irving, 1989), brute-force computation of  $d(\mu)$  is infeasible. Fortunately, by exploiting the structure of  $\mathcal{M}$ , we are able to efficiently reduce the computation of  $d(\mu)$  to a max-flow/min-cut problem of size quadratic in  $n$ . Thus, any number of efficient algorithms may be applied to compute  $d(\mu)$ .

### 6.1. The rotation poset and digraph

Our exposition follows the work of Gusfield (1987) and of Irving and Leather (1986) (see also Gusfield and Irving, 1989). Let  $\mu$  be a stable marriage. A **rotation  $\rho$  exposed by  $\mu$**  is a sequence of pairs  $(w_0, m_0), (w_1, m_1), \dots, (w_{r-1}, m_{r-1}) \in \mu$  such that for each  $i$ ,  $w_{i+1}$  is the first woman on  $m_i$ 's preference list that prefers  $m_i$  to her partner  $m_{i+1}$  in  $\mu$  (where addition is conducted modulo  $r$ ). Given  $\mu$  and  $\rho$ , we form a marriage called the **elimination of  $\rho$  from  $\mu$** ,

denoted  $\mu/\rho$ , which contains the pairs  $(w_1, m_0), (w_2, m_1), \dots, (w_0, m_{r-1})$ , in addition to all pairs from  $\mu$  that are not part of the rotation  $\rho$ . It is straightforward to verify that  $\mu/\rho$  is a stable marriage.

Let  $\mu_0$  denote the  $M$ -optimal stable marriage (the marriage found by the Gale-Shapley algorithm). Irving and Leather (1986) prove that every stable marriage can be obtained from  $\mu_0$  by successively eliminating a *unique* set of rotations that appear in  $\mu_0$  and subsequent stable marriages. Given a stable marriage  $\mu$ , let  $S_\mu$  denote this unique set of rotations, which can be eliminated (starting at  $\mu_0$ ) to obtain  $\mu$ . ( $S_\mu$  may contain some rotations that are not exposed in  $\mu_0$ , but only in subsequent marriages.)

We denote the set of all rotations exposed in one or more stable marriages in  $\mathcal{M}$  by  $\Pi(\mathcal{M})$ . We endow  $\Pi(\mathcal{M})$  with a partial order  $<$  where  $\rho < \sigma$  if  $\rho \in S_\mu$  for every  $\mu \in \mathcal{M}$  in which  $\sigma$  is exposed. In other words,  $\rho < \sigma$  if whenever  $\sigma$  is eliminated during the construction of a stable marriage by elimination from  $\mu_0$ , it is the case that  $\rho$  has been eliminated before  $\sigma$ . A subset  $S \subseteq \Pi(\mathcal{M})$  is (downward) **closed** if for all  $\sigma \in S$  and  $\rho < \sigma$ , we have that  $\rho \in S$ . Irving and Leather prove the following remarkable correspondence between closed subsets of  $\Pi(\mathcal{M})$  and stable marriages.

**Theorem 6.1** (Irving and Leather, 1986). *The map  $\mu \mapsto S_\mu$  is a bijection between  $\mathcal{M}$  and the set of closed subsets of  $\Pi(\mathcal{M})$ .*

For algorithmic purposes, it is advantageous to have a sparse representation of the partial order  $<$  on  $\Pi(\mathcal{M})$ , which preserves its closed subsets. To this end, Gusfield (1987) proved the following theorem. For the remainder of this section, we use the standard notation  $\tilde{O}$  to suppress  $\log n$  factors, which we find less interesting in the context of our discussion of time complexity in this section (in contrast to the discussion of communication and query complexity in the rest of this paper).

**Theorem 6.2** (Gusfield, 1987). *There exists a directed acyclic graph  $G(\mathcal{M})$  with vertices  $\Pi(\mathcal{M})$ , called the **rotation digraph**, whose transitive closure is the partial order  $<$  on  $\Pi(\mathcal{M})$ . (That is, a path from a vertex  $\rho$  to a vertex  $\sigma$  exists iff  $\rho < \sigma$ .)  $G(\mathcal{M})$  can be computed from the full preferences of all participants in time  $\tilde{O}(n^2)$ . In particular, the edge and vertex sets of  $G(\mathcal{M})$  both have cardinality  $\tilde{O}(n^2)$ .*

## 6.2. Rotation weights

In this section, we show how to assign weights to rotations  $\rho \in \Pi(\mathcal{M})$  in such a way that  $d(\mu, \mu')$  can be computed directly from  $d(\mu, \mu_0)$  and  $S_{\mu'}$ . Let  $\rho$  be any rotation and  $\mu'$  a stable marriage in which  $\rho$  is exposed. For any marriage  $\mu$ , we define the **weight**  $\gamma$  of  $\rho$  relative to  $\mu$  by

$$\gamma_\mu(\rho) = |\mu \cap (\mu'/\rho)| - |\mu \cap \mu'|.$$

(The absence of  $\mu'$  from the notation  $\gamma_\mu(\rho)$  becomes clear in Eq. (1) below.) That is,  $\gamma_\mu(\rho)$  is the net change, following the elimination of  $\rho$  from  $\mu'$ , in the number of pairs contained in the intersection of  $\mu$  and  $\mu'$ . We note that  $\gamma_\mu(\rho)$  can be computed directly from  $\rho$  (without being given an explicit stable marriage  $\mu'$  which exposes  $\rho$ ). Specifically, letting  $\rho' = \{(w_1, m_0), (w_2, m_1), \dots, (w_0, m_{r-1})\}$  be the set of pairs replacing  $\rho$  when eliminating  $\rho$  from any stable marriage,<sup>20</sup> we have

$$\gamma_\mu(\rho) = |\mu \cap \rho'| - |\mu \cap \rho|. \quad (1)$$

**Lemma 6.1.** *For any marriage  $\mu$  and stable marriage  $\mu' \in \mathcal{M}$ ,*

$$|\mu \cap \mu'| = |\mu \cap \mu_0| + \sum_{\rho \in S_{\mu'}} \gamma_\mu(\rho),$$

where  $\mu_0$  is the  $M$ -optimal stable marriage.

**Proof.** We argue by induction on  $|S_{\mu'}|$ . If  $S_{\mu'} = \emptyset$ , then  $\mu' = \mu_0$ , so the result is immediate. Suppose the claim is true for  $\mu'$  and  $\sigma$  is a rotation exposed in  $\mu'$ , then by the induction hypothesis and by definition of  $\gamma_\mu(\sigma)$ ,

$$|\mu \cap (\mu'/\sigma)| = |\mu \cap \mu'| + (|\mu \cap (\mu'/\sigma)| - |\mu \cap \mu'|) = |\mu \cap \mu_0| + \sum_{\rho \in S_{\mu'}} \gamma_\mu(\rho) + \gamma_\mu(\sigma),$$

which gives the desired result.  $\square$

Applying Lemma 6.1 and Theorem 6.1, we obtain the following result.

<sup>20</sup> Note that in general,  $\rho'$  is not a rotation.



**Theorem 6.3.** Let  $\mu$  be a marriage. Then

$$d(\mu) = d(\mu, \mu_0) - \max_S \sum_{\rho \in S} \gamma_\mu(\rho),$$

where the maximum is taken over closed subsets  $S \subseteq \Pi(\mathcal{M})$  (and where  $\mu_0$  is the  $M$ -optimal stable marriage).

**Proof.** By Lemma 6.1, for any stable marriage  $\mu'$ ,

$$d(\mu, \mu') = n - |\mu \cap \mu'| = n - |\mu \cap \mu_0| - \sum_{\rho \in S_{\mu'}} \gamma_\mu(\rho) = d(\mu, \mu_0) - \sum_{\rho \in S_{\mu'}} \gamma_\mu(\rho).$$

By Theorem 6.1,  $\mu' \mapsto S_{\mu'}$  is a bijection onto the set of closed subsets of  $\Pi(\mathcal{M})$ . Thus,

$$d(\mu) = \min_{\mu' \in \mathcal{M}} d(\mu, \mu') = d(\mu, \mu_0) - \max_S \sum_{\rho \in S} \gamma_\mu(\rho),$$

as desired.  $\square$

### 6.3. Reduction to max-flow/min-cut

We now wish to apply Theorems 6.2 and 6.3 to efficiently calculate  $d(\mu)$ . The divorce distance  $d(\mu, \mu_0)$  can easily be computed in  $\tilde{O}(n^2)$  time by using the Gale-Shapley algorithm to compute  $\mu_0$ . Since the partial order  $<$  on  $\Pi(\mathcal{M})$  is the transitive closure of  $G(\mathcal{M})$  (where  $G(\mathcal{M})$  is the rotation digraph described in Theorem 6.2), the closed subsets of  $<$  are precisely the (downward) **closed subsets** (vertex sets with no incoming edges) of  $G(\mathcal{M})$ , and so to compute  $d(\mu)$  it suffices to maximize  $\sum_{\rho \in S} \gamma_\mu(\rho)$  over closed subsets  $S$  of  $G(\mathcal{M})$ . Thus, we have reduced the problem of computing  $d(\mu)$  to finding a **maximum closed subset** (i.e. maximum-weight vertex set with no incoming edges) in a directed acyclic graph. This problem is well studied, in particular for its applications to open-pit mining (see, e.g. Picard, 1976; Hochbaum, 2001). For completeness and due to some differences in terminology between this paper and that of Picard (1976), we briefly describe the application to our specific maximum closed subset problem of Picard's efficient reduction (Picard, 1976) of maximum closed subset to max-flow/min-cut.<sup>21</sup>

Denote the vertex and edge sets of  $G(\mathcal{M})$  by  $V = \Pi(\mathcal{M})$  and  $E$  respectively. Let

$$V^+ = \{\rho \in V \mid \gamma_\mu(\rho) \geq 0\} \quad \text{and} \quad V^- = \{\rho \in V \mid \gamma_\mu(\rho) < 0\}.$$

We add a source vertex  $s$  and a sink vertex  $t$  to  $G(\mathcal{M})$  to form a new  $st$ -graph  $\tilde{G} = (\tilde{V}, \tilde{E})$  where  $\tilde{V} = V \cup \{s, t\}$  and

$$\tilde{E} = E \cup \{(s, \rho) \mid \rho \in V^-\} \cup \{(\rho, t) \mid \rho \in V^+\}.$$

We assign nonnegative capacity  $c(u, v)$  to each edge  $(u, v) \in \tilde{E}$  by

$$c(u, v) = \begin{cases} \infty & (u, v) \in E, \\ \gamma_\mu(\rho) & (u, v) = (\rho, t), \\ -\gamma_\mu(\rho) & (u, v) = (s, \rho). \end{cases}$$

**Theorem 6.4** (Picard, 1976). The sink set (i.e. the set of vertices on the same side as the sink  $t$ ) of a minimum  $st$ -cut in  $\tilde{G}$  is a maximum closed subset in  $G(\mathcal{M})$ .

In light of Theorem 6.4, we can reduce the computation of  $d(\mu)$  to known efficient algorithms for max-flow/min-cut. We summarize the procedure as follows.

**Theorem 6.5.** Given a marriage market  $(W, M, P_W, P_M)$  and a marriage  $\mu$ , running the procedure  $\text{DivorceDistance}(\mu, P_W, P_M)$  computes the divorce distance to stability  $d(\mu)$  in time  $\tilde{O}(n^4)$ .

We remark that since  $(P_W, P_M)$  has size  $\tilde{O}(n^2)$ , the runtime of  $\text{DivorceDistance}$  is nearly quadratic in the input size.

<sup>21</sup> Picard (1976) and Hochbaum (2001) use the term “closure”/“closed” to refer to an *upward* closed subset, i.e. a vertex set with no *outgoing* edges. This disagrees with Irving and Leather's (1986) (and our) usage of closed to mean *downward* closed (vertex set with no *incoming* edges), which appears to be more prevalent in the stable marriage literature. For consistency within this paper, when describing Picard's construction below and when phrasing Picard's theorem as Theorem 6.4, we perform the trivial needed modifications (flipping the direction of all edges as well as the roles of the source and sink vertices) to naturally present Picard's results with respect to downward rather than upward closed subsets.

---

**Algorithm 1**  $\text{DivorceDistance}(\mu, P_W, P_M)$  – compute  $d(\mu)$  with respect to preferences  $P_W, P_M$ .

---

1. Use the Gale-Shapley algorithm to compute  $\mu_0$  and compute  $d(\mu, \mu_0) = n - |\mu \cap \mu_0|$ .
  2. Construct the rotation digraph  $G(\mathcal{M})$  and the related graph  $\tilde{G}$ .
  3. Compute the weights in  $\tilde{G}$  by computing  $\gamma_\mu(\rho)$  for each rotation  $\rho \in G(\mathcal{M})$ , using Eq. (1).
  4. Find a minimum  $st$ -cut in  $\tilde{G}$ , and let  $S$  be the sink set in the cut.
  5. Return  $d(\mu, \mu_0) - \sum_{\rho \in S} \gamma_\mu(\rho)$ .
- 

**Proof of Theorem 6.5.** The correctness of  $\text{DivorceDistance}(\mu, P_W, P_M)$  follows immediately from Theorems 6.3 and 6.4. We analyze the runtime as follows. Step 1 can be computed in time  $\tilde{O}(n^2)$  using the Gale-Shapley algorithm and brute force computation of  $d(\mu, \mu_0)$ . For Step 2, by Theorem 6.2,  $G(\mathcal{M})$  (and hence  $\tilde{G}$ ) can also be computed in time  $\tilde{O}(n^2)$ . The weights in Step 3 can be computed in linear time for each rotation, so computing all the weights can be accomplished in time  $\tilde{O}(n^3)$ . For Step 4, the min-cut can be computed in time  $\tilde{O}(|\tilde{E}||\tilde{V}|) = \tilde{O}(n^4)$  using, for example, Hochbaum's algorithm (Hochbaum, 2001).  $\square$

We remark that since Algorithm 1 finds both  $\mu_0$  and  $S = S_{\mu'}$  for a stable marriage  $\mu'$  closest to  $\mu$ , it is a trivial task, which does not increase the asymptotic runtime complexity of Algorithm 1, to also compute  $\mu'$  in addition to computing  $d(\mu) = d(\mu, \mu')$ .

## 7. Commentary and open problems

A number of recent papers (Ashlagi et al., 2017; Gonczarowski, 2014) have touched on various aspects of the amorphic question of “how much do the preferences of the women in the Gale-Shapley algorithm affect the produced ( $M$ -optimal) stable marriage.” The fact that we prove our lower-bound result in a strong two-sided communication model (and not a weaker  $2n$ -sided communication model or an even-weaker query model) allows our results to also be viewed in the context of this line of research. Our communication lower bounds show that a significant amount of information about the preferences of the women is indeed needed in order to deduce the  $M$ -optimal stable marriage, as well as for solving any of the other problems described in Corollary 3.2.

One qualitative feature of the Gale-Shapley algorithm is that a single proposal at any point can precipitate a cascade of rejections that affects a large portion of the population. Thus, it is impossible for participants to know if their current partner is their final partner until the algorithm has terminated. Our results imply that this feature is common to *all* stable marriage mechanisms that dynamically refine a marriage and converge to a stable marriage. Indeed, consider any stable marriage algorithm and arbitrarily divide it into a “first” stage and a “second” stage. A consequence of Corollary 3.1 along with our novel definition of divorce distance is that if the query complexity of the first stage is significantly lower than that of querying the entire input, then after the first stage a large fraction of the participants may not yet be married to their final spouses.

In many real-world marriage markets, centralized clearinghouses are employed to prevent undesirable outcomes (Roth, 1984; Roth, 1991). Specifically, these clearinghouses were implemented to avoid “unraveling” – wherein participants are incentivized to match extremely early – as well as instability. While unraveling is undesirable in its own right, the early binding commitments made in an unraveling market have been shown to have adverse effects (Roth and Xing, 1994; Li and Rosen, 1998; Niederle and Roth, 2003; Fréchette et al., 2007), presumably because the early matches are necessarily made with incomplete knowledge about the market. A consequence of Corollary 3.2 (specifically, of Part (c)) is that any marriage mechanism that allows even a small fraction of participants to match in early binding commitments (i.e. before essentially all of the preferences are queried) cannot generally produce a stable marriage. Thus the empirical phenomenon of instability in decentralized markets where early-accepted proposals are binding, which was observed in Roth (1984, 1991), is not merely a feature of the particular marriage mechanisms that arose in practice, but is a general theoretical feature inherent in the stable marriage problem.

It is interesting to compare the  $\Omega(n^2)$  lower bound that is proved in this paper for the communication complexity of finding an approximately stable marriage to known complexity bounds for the problem of finding an approximately maximum-weight matching in a bipartite graph. Even though these two problems seem similar, the latter can be solved with  $\mathcal{O}(n \log n)$  communication (Dobzinski et al., 2014), i.e. with significantly less communication than many variants of the former. It is worthwhile to compare this surprising dissimilarity between these problems with a qualitatively similar message that emerges from a significantly different, recent line of work (Li, 2017; Ashlagi and Gonczarowski, 2018), which shows that finding a Pareto-efficient perfect matching requires considerably less strategic reasoning than finding a stable marriage.

The classic Gale-Shapley algorithm (Gale and Shapley, 1962) terminates after  $\mathcal{O}(n^2)$  steps, and each step consists of a message of  $\mathcal{O}(\log n)$  bits. Thus, the Gale-Shapley algorithm provides a communication upper bound of  $\mathcal{O}(n^2 \log n)$  for the problem of finding a stable marriage. As mentioned in the introduction, our Theorem 3.1 matches this up to a logarithmic factor, but it is not immediately clear how to close this gap.

**Open Problem 7.1.** What is the communication complexity of finding a stable marriage?

Our definition of  $(1 - \varepsilon)$ -stability is nonstandard. A more common notion of approximate stability is that a marriage induce few (say, at most  $\varepsilon n^2$ ) blocking pairs (see Eriksson and Häggström, 2008). As shown in Propositions 2.1, the blocking-pairs notion of approximate stability is strictly coarser than ours. It is therefore natural to ask if the  $\Omega(n^2)$  communication lower bound of Theorem 3.1 holds as well with respect to the more challenging concept of blocking-pairs approximate stability.

**Open Problem 7.2.** Is there a protocol  $\Pi$  that computes a marriage with at most  $\varepsilon n^2$  blocking pairs using  $o(n^2)$  communication?

Recently, Ostrovsky and Rosenbaum (2015) showed that it is possible to find a marriage with  $\varepsilon n^2$  blocking pairs for arbitrary  $\varepsilon > 0$  using  $\mathcal{O}(1)$  communication rounds for a distributed model of computation. While their result does not imply anything nontrivial about the total communication, we believe their techniques may be relevant for finding  $o(n^2)$  communication protocols for blocking-pairs approximate stability (if such protocols exist). Interestingly, an analogue of Theorem 3.2(a) (or of Corollary 3.2(a)) does not hold for blocking-pairs approximate stability.

**Theorem 7.1.** For every  $\varepsilon \geq \delta > 0$ , there exists a randomized communication protocol  $\Pi$  that determines whether a given marriage  $\mu$  induces at least  $\varepsilon n(n-1)$  blocking pairs or at most  $(\varepsilon - \delta)n(n-1)$  blocking pairs using  $\mathcal{O}(\log n)$  communication. In particular,  $\Pi$  determines whether  $\mu$  is stable or has  $\varepsilon n(n-1)$  blocking pairs using  $\mathcal{O}(\log n)$  communication.

**Proof sketch.** Choose an unmarried pair  $(w, m)$  uniformly at random from  $W \times M$ . If  $m$  prefers  $w$  over his spouse in  $\mu$ , the men query the women to see if  $w$  also prefers  $m$  over her spouse in  $\mu$  using  $\mathcal{O}(\log n)$  communication. The probability that  $(w, m)$  is a blocking pair is precisely  $\varepsilon'$ , where  $\varepsilon'$  is the fraction of blocking pairs in  $\mu$ . Repeat this procedure to estimate  $\varepsilon'$  to any desired accuracy in a bounded number of steps depending only on the desired accuracy.  $\square$

Theorem 3.2(c) shows that any protocol that produces a constant fraction of pairs in a stable marriage (regardless of which pairs are found) requires  $\Omega(n^2)$  communication. It would be interesting to improve this result (or find an efficient protocol) for finding even a single pair that appears in a stable marriage.

**Open Problem 7.3.** What is the communication complexity of finding a single pair  $(w, m)$  that appears in some/every stable marriage?

Finally, we notice that in contrast to Parts (a) and (c) of Theorem 3.2, our statement of Theorem 3.1 requires that  $\varepsilon < 1/2$ . It is natural to ask what can be obtained regarding other values of  $\varepsilon$ .

**Open Problem 7.4.** Fix  $\frac{1}{2} \leq \varepsilon < 1$ . What is the communication complexity of finding a  $(1 - \varepsilon)$ -stable marriage?

## Acknowledgments

Yannai Gonczarowski is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities. The work of Yannai Gonczarowski was supported in part by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [249159].

The work of Noam Nisan was supported in part by ISF grants 230/10 and 1435/14 administered by the Israeli Academy of Sciences, and by Israel-USA Bi-national Science Foundation (BSF) grant number 2014389.

The work of Rafail Ostrovsky was supported in part by NSF grants 09165174, 1065276, 1118126, 1136174 and 1619348; US-Israel BSF grant 2012366, The Okawa Foundation Research Award, IBM Faculty Research Award, Xerox Faculty Research Award, B. John Garrick Foundation Award, Teradata Research Award, and Lockheed-Martin Corporation Research Award. This material is based upon work supported in part by DARPA SafeWare program. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

We would like to thank the editor and the anonymous referees for many helpful comments.

## Appendix A. Asymptotic notation

Throughout this paper, we use standard computer-science asymptotic notation to describe the order of growth of single-dimensional functions of natural numbers. For example, for a positive function  $f$ , we write  $f(n) = \mathcal{O}(g(n))$  where  $g$  is also a positive function (usually, one simple to write down), if there exist positive numbers  $\overline{M}$  and  $N$  such that  $f(n) \leq \overline{M} \cdot g(n)$  for all  $n > N$ . (Equivalently and succinctly, we write  $f = \mathcal{O}(g(n))$  if  $\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$ .) Intuitively, one may find it helpful to read the notation  $f = \mathcal{O}(g(n))$  as  $f \in \mathcal{O}(g)$ , where  $\mathcal{O}(g)$  is the class of functions whose order of growth (as  $n$  grows

large) is at most that of  $g$ . This notation allows for the simplification of the exposition of many results, where the order of magnitude of the result serves as the main message. For example, for a highly complex function  $f$  that satisfies  $f = \mathcal{O}(n^2)$  (that is, does not grow any faster than  $g$ , where  $g(n) = n^2$  for all  $n$ ), instead of specifying  $f$  and saying that some algorithm takes at most  $f$  steps, one could concisely say that this algorithm takes at most  $\mathcal{O}(n^2)$  steps, without the need to explicitly write down the complex function  $f$ . The following table summarizes this and other similar standard “order of growth” notation used throughout this paper.

Notation	Explicit definition	Succinct definition	Informal meaning
$f = \mathcal{O}(g(n))$	$\exists \overline{M}, N : \forall n > N : f(n) \leq \overline{M} \cdot g(n)$	$\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$	$f$ grows at most as fast as $g$
$f = \Omega(g(n))$	$\exists \underline{M} > 0, N : \forall n > N : f(n) \geq \underline{M} \cdot g(n)$	$\liminf_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0$	$f$ grows at least as fast as $g$
$f = \Theta(g(n))$	$\exists \overline{M} > 0, \underline{M}, N : \forall n > N : \underline{M} \cdot g(n) \leq f(n) \leq \overline{M} \cdot g(n)$	$f = \mathcal{O}(g(n))$ & $f = \Omega(g(n))$	$f$ grows as fast as $g$
$f = o(g(n))$	$\forall \overline{M} > 0 : \exists N : \forall n > N : f(n) \leq \overline{M} \cdot g(n)$	$\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$	$f$ grows slower than $g$
$f = \omega(g(n))$	$\forall \underline{M} : \exists N : \forall n > N : f(n) \geq \underline{M} \cdot g(n)$	$\liminf_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty$	$f$ grows faster than $g$

## Appendix B. Embedding arbitrary preferences into complete preferences

This section contains the remaining technical details needed to complete the direct proof that is given in Section 4 of our lower bound on the communication complexity of finding an exactly stable marriage.

**Definition B.1** (Submarriage). Let  $W'$  and  $M'$  be disjoint sets. A marriage  $\mu$ , between a subset  $W$  of  $W'$  and a subset  $M$  of  $M'$ , is said to be a **submarriage** of a marriage  $\mu'$  between  $W'$  and  $M'$ , if for every  $w \in W$  and  $m \in M$ , we have  $\mu'(w) = m$  iff  $\mu(w) = m$ .

**Lemma B.1** (Finding a stable marriage (arbitrary preferences)  $\hookleftrightarrow$  Finding a stable marriage (full preferences)). Let  $n \in \mathbb{N}$ , and let  $W, W', M$  and  $M'$  be pairwise-disjoint sets, each of cardinality  $n$ . There exist functions  $P_{W \cup W'} : \mathcal{P}(W, M) \rightarrow \mathcal{F}(W \cup W', M \cup M')$  and  $P_{M \cup M'} : \mathcal{P}(M, W) \rightarrow \mathcal{F}(M \cup M', W \cup W')$  such that for every  $P_W \in \mathcal{P}(W, M)$  and  $P_M \in \mathcal{P}(M, W)$ , and for every (possibly imperfect) marriage  $\mu$  between  $W$  and  $M$ , the following are equivalent.

- $\mu$  is stable with respect to  $P_W$  and  $P_M$ .
- $\mu$  is a submarriage of some marriage between  $W \cup W'$  and  $M \cup M'$  that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .

**Proof.** <sup>22</sup> Denote  $W = \{w_1, \dots, w_n\}$ ,  $M = \{m_1, \dots, m_n\}$ ,  $W' = \{w'_1, \dots, w'_n\}$ , and  $M' = \{m'_1, \dots, m'_n\}$ .

To define  $P_{W \cup W'}(P_W)$ , for every  $i$  we define the preference list of  $w_i$  to consist of her preference list in  $P_W$  (in the same order), followed by  $m'_i$ , followed by all other men in arbitrary order; we define the preference list of  $w'_i$  to consist of  $m_i$ , followed by all other men in arbitrary order. Similarly, to define  $P_{M \cup M'}(P_M)$ , for every  $j$  we define the preference list of  $m_j$  to consist of his preference list in  $P_M$  (in the same order), followed by  $w'_j$ , followed by all other women in arbitrary order; we define the preference list of  $m'_j$  to consist of  $w_j$ , followed by all other women in arbitrary order.

It is straightforward to verify that the lemma holds with respect to these definitions of  $P_{W \cup W'}$  and  $P_{M \cup M'}$ ; the details are left to the reader.  $\square$

**Remark B.1.** It is straightforward to embed the problem of finding a stable marriage with respect to full preference lists in that of finding a stable marriage with respect to arbitrary preference lists, as the former is a special case of the latter.

## Appendix C. Determining the marital status of a given couple or participant

In this appendix, we give an alternate proof of Part (b) of Theorem 3.2 that uses the construction of Section 4. We prove Theorem 3.2(b) once again using Theorem 2.4, by embedding disjointness (without assuming unique intersection) in both problems. We embed disjointness via an intermediate problem of determining whether a given participant is single (i.e.

<sup>22</sup> Our construction in this proof is essentially a one-to-one version of the many-to-many construction given in Corollary 31 of Gonczarowski and Friedgut (2013).

not married to anyone) in some stable marriage, given profiles of arbitrary (i.e. not necessarily full) preference lists.<sup>23</sup> We therefore obtain the same lower bounds for this natural problem as well.

**Lemma C.1** (*Disjointness  $\leftrightarrow$  Is participant single?*). Let  $n \in \mathbb{N}$ , let  $W$  and  $M$  be disjoint sets such that  $|W| = |M| = 2n$ , and let  $p \in W \cup M$ . There exist functions  $P_W : \{0, 1\}^{[n]^2} \rightarrow \mathcal{P}(W, M)$  and  $P_M : \{0, 1\}^{[n]^2} \rightarrow \mathcal{P}(M, W)$  such that for every  $\bar{x} = (x_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$  and  $\bar{y} = (y_j^i)_{(i,j) \in [n]^2} \in \{0, 1\}^{[n]^2}$ , the following are equivalent.

- $p$  is single in some stable marriage with respect to  $P_W(\bar{x})$  and  $P_M(\bar{y})$ .
- $\text{DISJ}(\bar{x}, \bar{y}) \neq 0$ .

**Proof.** Assume without loss of generality that  $p \in W$  and denote  $w = p$ . Denote  $W = \{w_1, \dots, w_n, w, w'_2, w'_3, \dots, w'_n\}$  and  $M = \{m_1, \dots, m_n, m'_1, \dots, m'_n\}$ .

To define  $P_W(\bar{x})$ , for every  $i$  we define the preference list of  $w_i$  to consist of all  $m_j$  such that  $x_j^i = 1$ , in arbitrary order (say, sorted by  $j$ ), followed by  $m'_i$  (with all other men absent). We define the preference list of  $w$  to consist of all  $m'_j$ , in arbitrary order (say, sorted by  $j$ ), with all other men absent. We define the preference list of every  $w'_i$  to be empty (these women can be ignored, and are defined purely for aesthetic reasons – so that  $W$  and  $M$  be of equal cardinality). To define  $P_M(\bar{y})$ , for every  $j$  we define the preference list of  $m_j$  to consist of all  $w_i$  such that  $y_j^i = 1$ , in arbitrary order (say, sorted by  $i$ ), with all other women absent. For every  $j$  we define the preference list of  $m'_j$  to consist of  $w_j$ , followed by  $w$  (with all other women absent).

Let  $\mu'_{\text{id}}$  be the marriage in which  $w_i$  is married to  $m'_i$  for every  $i$ , and in which all other participants are single. We first show that  $\text{DISJ}(\bar{x}, \bar{y}) \neq 0$  iff  $\mu'_{\text{id}}$  is stable, and then show that  $\mu'_{\text{id}}$  is stable iff  $w = p$  is single in some stable marriage; we commence with the former.

We begin by noting that every participant that is married in  $\mu'_{\text{id}}$  is married to someone on their preference list; therefore,  $\mu'_{\text{id}}$  is stable iff no pair would rather deviate. Obviously, no  $w'_i$  would rather deviate with anyone. Furthermore, while  $w$  would rather deviate with any  $m'_j$ , these are all married to their top choices, and so none of them would deviate with  $w$ . Since for every  $i$ , the preference list of  $w_i$  consists of  $m'_i$  and of a subset of  $\{m_j\}_{j \neq i}$ , we therefore have that  $\mu'_{\text{id}}$  is unstable iff there exists  $(i, j) \in [n]^2$  such that both  $m_j \succ_{w_i} m'_i$  and  $w_i$  is on the preference list of  $m_j$ . Similarly to the proof of Lemma 4.2, this holds precisely if there exists  $(i, j) \in [n]^2$  such that  $x_j^i = 1$  and  $y_j^i = 1$ , which holds iff  $\text{DISJ}(\bar{x}, \bar{y}) = 0$ .

We complete the proof by showing that  $\mu'_{\text{id}}$  is stable iff  $w = p$  is single in some stable marriage. The first direction follows immediately from the fact that  $w$  is single in  $\mu'_{\text{id}}$ . For the second direction, assume that there exists a stable marriage  $\mu$  in which  $w$  is single. By stability of  $\mu$  and since all men on the preference list of  $w$  have  $w$  on their preference list, all such men are married in  $\mu$  and prefer their spouses over  $w$ . Therefore, for every  $j$ , we have that  $m'_j$  is married to  $w_j$  in  $\mu$ . By stability of  $\mu$ , every  $w'_i$  is single in  $\mu$ . As  $\mu$  and  $\mu'_{\text{id}}$  coincide on all women, we have that  $\mu = \mu'_{\text{id}}$ . Therefore,  $\mu'_{\text{id}} = \mu$  is stable and the proof is complete.  $\square$

**Corollary C.1** (*Complexity of determining the marital status of a given participant*). The lower bounds of Theorem 3.2 and Corollary 3.2 apply also to the problem of determining whether a given participant  $p \in W \cup M$  is single in some (equivalently, in every) stable marriage, where  $P_W \in \mathcal{P}(W, M)$  and  $P_M \in \mathcal{P}(M, W)$  are arbitrary (i.e. not necessarily full) preference lists.

**Lemma C.2** (*Is participant single?  $\leftrightarrow$  Is couple sometimes/always married?*). Let  $n \in \mathbb{N}$ , and let  $W, W', M$  and  $M'$  be pairwise-disjoint sets, each of cardinality  $n$ ; let  $w \in W$  and  $m' \in M'$ . There exist functions  $P_{W \cup W'} : \mathcal{P}(W, M) \rightarrow \mathcal{F}(W \cup W', M \cup M')$  and  $P_{M \cup M'} : \mathcal{P}(M, W) \rightarrow \mathcal{F}(M \cup M', W \cup W')$  such that for every  $P_W \in \mathcal{P}(W, M)$  and  $P_M \in \mathcal{P}(M, W)$ , the following are equivalent.

- $w$  is single in some marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M$ .
- $w$  and  $m'$  are married in **some** marriage between  $W \cup W'$  and  $M \cup M'$  that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .
- $w$  and  $m'$  are married in **every** marriage between  $W \cup W'$  and  $M \cup M'$  that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .

**Proof.** The proof is similar to that of Lemma B.1. Denote  $W = \{w_1 = w, w_2, \dots, w_n\}$ ,  $M = \{m_1, \dots, m_n\}$ ,  $W' = \{w'_1, \dots, w'_n\}$ , and  $M' = \{m'_1 = m', m'_2, \dots, m'_n\}$ .

To define  $P_{W \cup W'}(P_W)$ , for every  $i$  we define the preference list of  $w_i$  to consist of her preference list in  $P_W$  (in the same order), followed by  $m'_i$ , followed by all other men in arbitrary order; we define the preference list of  $w'_i$  to consist of  $m_i$ , followed by all other men in arbitrary order. Similarly, to define  $P_{M \cup M'}(P_M)$ , for every  $j$  we define the preference list

<sup>23</sup> By Theorem 2.3 (in conjunction with 2.1), this is equivalent to whether this participant is single in *every* stable marriage.



of  $m_j$  to consist of his preference list in  $P_M$  (in the same order), followed by  $w'_j$ , followed by all other women in arbitrary order; we define the preference list of  $m'_j$  to consist of  $w_j$ , followed by all other women in arbitrary order.

Similarly to the proof of Lemma B.1, we have that  $w$  is single in some marriage  $\mu$  between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M$  iff  $w$  and  $m'$  are married in some marriage (a corresponding “supermarriage” of  $\mu$ ) between  $W \cup W'$  and  $M \cup M'$  that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ . Additionally, by Theorem 2.3 (in conjunction with Theorem 2.1), we have:  $w$  is single in some marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M \iff w$  is single in every marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M \iff w$  and  $m'$  are married in every marriage between  $W \cup W'$  and  $M \cup M'$  that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .  $\square$

#### Appendix D. Verifying the output of a given stable marriage mechanism

As noted in Section 3, while our lower bound for verifying the stability of a given marriage is tight, we do now know whether our lower bound for finding a stable marriage is tight as well. We note that we do not even know a tight lower bound for verifying whether a given marriage is the  $M$ -optimal stable marriage.

**Open Problem D.1.** What is the worst-case complexity of verifying whether a given marriage is the  $M$ -optimal stable marriage?

As in the case of Open Problem 1.1, we do not have any  $o(n^2 \log n)$  algorithm for verification of the  $M$ -optimal stable marriage, even randomized and even in the strong two-party communication model, nor do we have any  $\omega(n^2)$  lower bound, even for deterministic algorithms and even in the simple comparison model.

In this section, we derive a  $\Omega(n^2)$  lower bound for verification of the  $M$ -optimal stable marriage. In fact, we show this lower bound not only for verifying the  $M$ -optimal stable marriage, but also for verifying the output of any other stable marriage mechanism.

**Definition D.1** (*Stable marriage mechanism*). Let  $n \in \mathbb{N}$ , let  $W$  and  $M$  be disjoint sets such that  $|W| = |M| = n$ . A **stable marriage mechanism** is a function  $f$  from  $\mathcal{F}(W, M) \times \mathcal{F}(M, W)$  to the set of perfect marriages between  $W$  and  $M$ , such that for every  $P_W \in \mathcal{F}(W, M)$  and  $P_M \in \mathcal{F}(M, W)$ , the marriage  $f(P_W, P_M)$  is stable with respect to  $P_W$  and  $P_M$ .

**Example D.1** ( *$M$ -optimal stable marriage mechanism*). The function  $f_{M\text{-Opt}}$ , defined such that  $f_{M\text{-Opt}}(P_W, P_M)$  is the  $M$ -optimal stable marriage with respect to  $P_W$  and  $P_M$ , is a well-defined stable marriage mechanism by Theorem 2.1.

**Corollary D.1** (*Complexity of computing the output of a given stable marriage mechanism*). By Theorem 3.1, we have that for every stable marriage mechanism  $f$ , the worst-case randomized query complexity (as defined in Corollary 3.1) as well as the worst-case communication complexity of computing  $f$  is  $\Omega(n^2)$ .

**Theorem D.1** (*Complexity of verifying the output of a given stable marriage mechanism*). Let  $n \in \mathbb{N}$ , let  $W$  and  $M$  be disjoint sets such that  $|W| = |M| = n$ , fix a stable marriage mechanism  $f$  and let  $P_W \in \mathcal{F}(W, M)$  and  $P_M \in \mathcal{F}(M, W)$ . Let  $\mu_{\text{id}}$  be the perfect marriage in which  $w_i$  is married to  $m_i$  for every  $i$ . The worst-case randomized query complexity, as well as the worst-case randomized communication complexity, of determining whether  $f(P_W, P_M) = \mu_{\text{id}}$  is  $\Omega(n^2)$ .

Theorem D.1 may be proven either via a direct application of the machinery of Section 5, or using the machinery of Section 4, with Lemma B.1 replaced by the following lemma.

**Lemma D.1.** Let  $n \in \mathbb{N}$ , and let  $W = \{w_1, \dots, w_n\}$ ,  $M = \{m_1, \dots, m_n\}$ ,  $W' = \{w'_1, \dots, w'_n\}$  and  $M' = \{m'_1, \dots, m'_n\}$  be pairwise-disjoint sets, each of cardinality  $n$ . Let  $\mu_{\text{id}}$  be the perfect marriage between  $W$  and  $M$  in which  $w_i$  is married to  $m_i$  for every  $i$ , and let  $\mu'_{\text{id}}$  be the perfect marriage between  $W \cup W'$  and  $M \cup M'$  in which for every  $i$ , both  $w_i$  is married to  $m_i$  and  $w'_i$  is married to  $m'_i$ . There exist functions  $P_{W \cup W'} : \mathcal{P}(W, M) \rightarrow \mathcal{F}(W \cup W', M \cup M')$  and  $P_{M \cup M'} : \mathcal{P}(M, W) \rightarrow \mathcal{F}(M \cup M', W \cup W')$  such that for every  $P_W \in \mathcal{P}(W, M)$  and  $P_M \in \mathcal{P}(M, W)$ , both of the following hold.

- (a) If  $\mu_{\text{id}}$  is the unique stable marriage with respect to  $P_W$  and  $P_M$ , then  $\mu'_{\text{id}}$  is the unique stable marriage with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .
- (b) If  $\mu_{\text{id}}$  is unstable with respect to  $P_W$  and  $P_M$ , then  $\mu'_{\text{id}}$  is unstable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ .

**Proof.** We define  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$  as in Lemma B.1, only with  $M'$  appearing sorted by  $j$  (as opposed to in arbitrary order) on the preference lists of  $W'$ , and with  $W'$  appearing sorted by  $i$  (as opposed to in arbitrary order) on the preference lists of  $M'$ . By Lemma B.1, we have both that (b) holds, and that if  $\mu_{\text{id}}$  is the unique stable marriage with respect to  $P_W$  and  $P_M$ , then it is a submarriage of every marriage that is stable with respect to  $P_{W \cup W'}(P_W)$  and  $P_{M \cup M'}(P_M)$ ; it is straightforward to show that every “supermarriage” of  $\mu_{\text{id}}$ , apart from  $\mu'_{\text{id}}$ , is unstable, thus proving (a) as well.  $\square$

**Open Problem D.2.** Is there a stable marriage mechanism whose worst-case output verification complexity is  $\Theta(n^2)$ ? Which stable marriage mechanisms have the lowest asymptotic worst-case output verification complexity?

## Appendix E. Nondeterminism

All the lower bounds in this paper are based upon reductions to the well-studied communication complexity of the disjointness function. Since the disjointness function also has  $\Theta(n)$  **nondeterministic** communication complexity (Kushilevitz and Nisan, 1997), it follows that all our lower bounds apply not only to randomized communication complexity, but also to nondeterministic communication complexity. For nondeterministic communication complexity, the  $\Omega(n^2)$  lower bound for finding a stable marriage is in fact tight (and so still is the  $\Omega(n^2)$  bound for verification of stability).

For the decision problem of verifying the stability of a given marriage, the **co**-nondeterministic communication complexity may be easily seen to be  $\Theta(\log n)$ . In contrast, we note that the proof of Theorem 3.2(b) may be easily adapted to show a  $\Omega(n^2)$  lower bound also for the co-nondeterministic communication complexities of determining the marital status of a given couple.

**Theorem E.1** (Nondeterministic communication complexity of determining the marital status of a given couple). *In the notation of Theorem 3.2(b), both the nondeterministic and co-nondeterministic communication complexities of determining whether  $w$  and  $m$  are married in some/every stable marriage are  $\Omega(n^2)$ .*

For completeness, we show this lower bound also for the nondeterministic and co-nondeterministic communication complexities of the intermediate problem of determining whether a given participant is single, which we presented in Appendix C. (This proof also yields Theorem E.1 using the tools of that appendix and of Section 4.) These lower bounds follow from the results of Appendix C in conjunction with the following lemma.

**Lemma E.1** (Is participant single?  $\leftrightarrow$   $\neg$  Is participant single?). *Let  $n \in \mathbb{N}$ , let  $W$  and  $M$  be sets such that  $|W| = |M| = n$ , and let  $w'$  and  $m'$  such that  $W$ ,  $M$ ,  $\{w'\}$  and  $\{m'\}$  are pairwise disjoint; let  $w \in W$ . There exist functions  $P_{W \cup \{w'\}} : \mathcal{P}(W, M) \rightarrow \mathcal{P}(W \cup \{w'\}, M \cup \{m'\})$  and  $P_{M \cup \{m'\}} : \mathcal{P}(M, W) \rightarrow \mathcal{P}(M \cup \{m'\}, W \cup \{w'\})$  such that for every  $P_W \in \mathcal{P}(W, M)$  and  $P_M \in \mathcal{P}(M, W)$ , the following are equivalent.*

- $w$  is single in some marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M$ .
- $m'$  is married in every marriage between  $W \cup \{w'\}$  and  $M \cup \{m'\}$  that is stable with respect to  $P_{W \cup \{w'\}}(P_W)$  and  $P_{M \cup \{m'\}}(P_M)$ .

**Proof.** To define  $P_{W \cup \{w'\}}(P_W)$ , we define the preference list of  $w$  as her preference list in  $P_W$  (in the same order), followed by  $m'$ ; we define the preference list of every other woman in  $W$  as her preference list in  $P_W$  (in the same order and with  $m'$  absent), and define the preference list of  $w'$  to be empty (once again,  $w'$  can be ignored, and is defined purely for aesthetic reasons – so that  $W \cup \{w'\}$  and  $M \cup \{m'\}$  be of equal cardinality). To define  $P_{M \cup \{m'\}}(P_M)$ , we define the preference list of every man in  $M$  as his preference list in  $P_M$  (in the same order and with  $w'$  absent); we define the preference list of  $m'$  to consist solely of  $w$ .

Directly from definition of  $P_{M \cup \{m'\}}$  and  $P_{W \cup \{w'\}}$ , we have that a natural bijection  $\mu \mapsto \mu'$  exists between stable marriages with respect to  $P_W$  and  $P_M$  and stable marriages with respect to  $P_{W \cup \{w'\}}(P_W)$  and  $P_{M \cup \{m'\}}(P_M)$ ; this bijection is given by:

- If  $w$  is married in  $\mu$ , then  $\mu' = \mu$  (with  $m'$  and  $w'$  single in  $\mu'$ ).
- If  $w$  is single in  $\mu$ , then  $\mu'$  is the marriage obtained from  $\mu$  by marrying  $w$  to  $m'$  (with  $w'$  once again single in  $\mu'$ ).

Once again by Theorem 2.3 (in conjunction with Theorem 2.1), and by the existence of this bijection, we have:  $w$  is single in some marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M \iff w$  is single in every marriage between  $W$  and  $M$  that is stable with respect to  $P_W$  and  $P_M \iff m'$  is married in every marriage between  $W \cup \{w'\}$  and  $M \cup \{m'\}$  that is stable with respect to  $P_{W \cup \{w'\}}(P_W)$  and  $P_{M \cup \{m'\}}(P_M)$ .  $\square$

We note that the nondeterministic lower bound of  $\Omega(n^2)$  for determining whether a given couple is married in some stable marriage, as well as the co-nondeterministic lower bound of  $\Omega(n^2)$  for determining whether a given couple is married in every stable marriage (and both the nondeterministic and co-nondeterministic lower bounds of  $\Omega(n^2)$  for determining whether a given participant is single in some/every stable marriage), is in fact tight. (Recall that we do not know whether any of these problems can be deterministically or even probabilistically solved using  $o(n^2 \log n)$  communication.) The questions of a tight co-nondeterministic lower bound for the former problem and a tight nondeterministic lower bound for the latter remain open in all query models. We note that the latter problem may be solved by checking whether the pair in question is married in both the  $M$ -optimal stable marriage and the  $W$ -optimal stable marriage; a  $\mathcal{O}(n^2)$ -Boolean-queries algorithm (even a nondeterministic one) for verification of the  $M$ -optimal stable marriage (see Open Problem D.1 in Appendix D) would therefore also settle the question of the nondeterministic communication complexity of this problem.

## Appendix F. Optimality of deferred acceptance with respect to queries onto women

Gale and Shapley's proof (Gale and Shapley, 1962) of Theorem 2.1 is constructive, providing an efficient algorithm for finding the  $M$ -optimal stable marriage. In this algorithm, men are asked queries of the form “which woman is next on the preference list of man  $m$  after woman  $w$ ?” (or alternatively, “which woman does man  $m$  rank at place  $k$ ?”), while women are asked queries of the form “whom does woman  $w$  prefer most out of the set of men  $\tilde{M}$ ?”; all of these queries require an answer of length  $\mathcal{O}(\log n)$  bits.

Dubins and Freedman (1981) presented a variant of Gale and Shapley's algorithm, which runs in the same worst-case time complexity, but performs a significantly more limited class of queries, namely only pairwise-comparison queries, onto women. In Open Problem 1.1 in the Introduction, we raise the question of a tight lower bound for the complexity of finding a stable marriage using only such queries for both women and men. In this section, we show that regardless of how complex the queries onto the men may be, no algorithm for finding any stable marriage (and even no algorithm for verifying the stability of a given marriage, when input a stable marriage) that performs only pairwise-comparison queries onto women, may perform any less such queries onto them than Dubins and Freedman's variant of Gale and Shapley's algorithm (given the same preference lists). For the duration of this section, let  $n \in \mathbb{N}$ , let  $W$  and  $M$  be disjoint sets such that  $|W| = |M| = n$ .

**Definition F.1** (Pairwise-comparison query). A **pairwise-comparison query** onto  $W$  is a query of whether  $m \succ_w m'$  for some given  $w \in W$  and  $m, m' \in M$ .

**Definition F.2** (Men-proposing deferred-acceptance algorithm; Dubins and Freedman, 1981). The following algorithm is henceforth referred to as the **men-proposing deferred-acceptance algorithm**: The algorithm is initialized with all women and all men being **provisionally single**, and concludes when no man is provisionally single. The algorithm is divided into steps, to which we refer as **nights**. On each night, an arbitrary provisionally-single man  $m$  is chosen, and serenades under the window of the woman  $w$  ranked highest on his preference list among those who have not (yet) rejected him. If  $w$  is provisionally single, then  $m$  and  $w$  are **provisionally married** to each other. Otherwise, i.e. if  $w$  is already provisionally married to some man  $m'$ , then if  $m \succ_w m'$ , then  $w$  rejects  $m'$ , who becomes provisionally single, and  $w$  and  $m$  are provisionally married to each other; otherwise,  $w$  rejects  $m$ , who remains provisionally single. The algorithm stops when no provisionally-single men remain, and the couples married by the output marriage are exactly those that are provisionally married when the algorithm stops.

**Theorem F.1** (Dubins and Freedman, 1981). Let  $P_W \in \mathcal{F}(W, M)$  and  $P_M \in \mathcal{F}(M, W)$  be profiles of full preference lists for  $W$  over  $M$  and for  $M$  over  $W$ , respectively. The men-proposing deferred-acceptance algorithm stops after  $\mathcal{O}(n^2)$  nights, and yields the  $M$ -optimal stable marriage.

**Remark F.1.** Let  $P_W \in \mathcal{F}(W, M)$  and let  $P_M \in \mathcal{F}(M, W)$ . All runs of the men-proposing deferred-acceptance algorithm (given  $P_W$  and  $P_M$ ) perform the same number of pairwise-comparison queries onto  $W$ .

**Theorem F.2** (Optimality of men-proposing deferred-acceptance algorithm with respect to pairwise-comparison queries onto  $W$ ). For any profiles  $P_W \in \mathcal{F}(W, M)$  and  $P_M \in \mathcal{F}(M, W)$  of full preference lists for  $W$  over  $M$  and for  $M$  over  $W$ , respectively, every algorithm for finding or verifying a stable marriage (for the latter – when input any marriage that is stable with respect to  $P_W$  and  $P_M$ ) that only performs pairwise-comparison queries onto  $W$  (and arbitrary queries onto  $M$ ), performs no less queries onto  $W$  than the men-proposing deferred-acceptance algorithm, when input  $P_W$  and  $P_M$ .

**Remark F.2.** An analogous result may similarly be shown to hold with respect to profiles of arbitrary preference lists, and finding/verifying a possibly-imperfect stable marriage.

**Definition F.3.** Let  $\mu$  be a perfect marriage between  $W$  and  $M$ . By slight abuse of notation, we denote the woman married to a man  $m \in M$  in  $\mu$  by  $\mu(m)$  instead of  $\mu^{-1}(m)$ .

**Proof of Theorem F.2.** Let  $A$  be a run of the men-proposing deferred-acceptance algorithm with respect to  $P_W$  and  $P_M$ , and let  $B$  be a given run of an algorithm for finding/verifying a stable marriage with respect to  $P_W$  and  $P_M$ . Let  $Q \subseteq W \times M^2$  be the set of triples  $(w, m, m')$  such that either the query of whether  $m \succ_w m'$  was performed onto  $W$  during  $B$  and answered positively, or the query of whether  $m' \succ_w m$  was performed onto  $W$  during  $B$  and answered negatively. By definition, at least  $|Q|$  queries onto  $W$  are performed during  $B$ . Let  $\mu$  be the  $M$ -optimal stable marriage with respect to  $P_W$  and  $P_M$ , i.e. the marriage output by  $A$ . Let  $R = \{(w, m) \mid w \text{ rejects } m \text{ during } A\} \subseteq W \times M$ . By definition, we note that the number of queries onto  $W$  during  $A$  equals the number of rejections performed during  $A$ , and so, as no woman rejects the same man twice, equals  $|R|$ . It is therefore enough to show that  $|R| \leq |Q|$  in order to complete the proof.

Let  $\mu'$  be the output of  $B$  if it is a run of an algorithm for finding a stable marriage, or the input to  $B$  if it is a run of an algorithm for verifying stability; either way,  $\mu'$  a stable marriage with respect to  $P_W$  and  $P_M$ . We claim that  $w \succ_m \mu'(m)$  for every  $(w, m) \in R$ . Indeed, as  $m$  serenades under women's windows during  $A$  in descending order of preference, the

fact that  $w$  rejects  $m$  during  $A$  implies  $w \succ_m \mu(m)$ . By Theorem F.1, we thus have  $w \succ_m \mu(m) \succeq_m \mu'(m)$ , as claimed. As  $B$  guarantees the stability of  $\mu'$ , it must therefore ascertain that  $\mu'(w) \succ_w m$  for every  $(w, m) \in R$ ; therefore, as only pairwise-comparison queries are performed onto  $W$  during  $B$ , there exists  $m' \in M$  such that  $(w, m', m) \in Q$ . We have thus shown that  $R$  is contained in the projection of  $Q$  over its first and last coordinates, and therefore  $|R| \leq |Q|$ , and the proof is complete.  $\square$

## References

- Ashlagi, I., Gonczarowski, Y.A., 2018. Stable matching mechanisms are not obviously strategy-proof. *J. Econ. Theory* 177, 405–425.
- Ashlagi, I., Kanoria, Y., Leshno, J.D., 2017. Unbalanced random matching markets: the stark effect of competition. *J. Polit. Econ.* 125 (1), 69–98.
- Bei, X., Chen, N., Zhang, S., 2013. On the complexity of trial and error. In: *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 31–40.
- Bhatnagar, N., Greenberg, S., Randall, D., 2008. Sampling stable marriages: why spouse-swapping won't work. In: *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1223–1232.
- Bogomolnaia, A., Laslier, J.-F., 2007. Euclidean preferences. *J. Math. Econ.* 43 (2), 87–98.
- Chou, J.-H., Lu, C.-J., 2010. Communication requirements for stable marriages. In: *Proceedings of the 7th International Conference on Algorithms and Complexity (CIAC)*, pp. 371–382.
- Dobzinski, S., Nisan, N., Oren, S., 2014. Economic efficiency requires interaction. In: *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 233–242, full version available at arXiv:1311.4721.
- Dubins, L.E., Freedman, D.A., 1981. Machiavelli and the Gale-Shapley algorithm. *Am. Math. Mon.* 88 (7), 485–494.
- Eriksson, K., Häggström, O., 2008. Instability of matchings in decentralized markets with various preference structures. *Int. J. Game Theory* 36 (3), 409–420.
- Fréchette, G.R., Roth, A.E., Ünver, M.U., 2007. Unraveling yields inefficient matchings: evidence from post-season college football bowls. *Rand. J. Econ.* 38 (4), 967–982.
- Gale, D., Shapley, L.S., 1962. College admissions and the stability of marriage. *Am. Math. Mon.* 69 (1), 9–15.
- Gonczarowski, Y.A., 2014. Manipulation of stable matchings using minimal blacklists. In: *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, p. 449.
- Gonczarowski, Y.A., Friedgut, E., 2013. Sisterhood in the Gale-Shapley matching algorithm. *Electron. J. Comb.* 20 (2), #P12, 18 pp.
- Gonczarowski, Y.A., Nisan, N., Ostrovsky, R., Rosenbaum, W., 2015. A stable marriage requires communication. In: *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1003–1017.
- Gusfield, D., 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16 (1), 111–128.
- Gusfield, D., Irving, R.W., 1989. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press.
- Hochbaum, D.S., 2001. A new-old algorithm for minimum-cut and maximum-flow in closure graphs. *Networks* 37 (4), 171–193.
- Irving, R.W., Leather, P., 1986. The complexity of counting stable marriages. *SIAM J. Comput.* 15 (3), 655–667.
- Kalyanasundaram, B., Schintger, G., 1992. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.* 5 (4), 545–557.
- Knuth, D.E., 1976. *Marriage stables et leurs relations avec d'autres problèmes combinatoires*. Les Presses de l'Université de Montréal.
- Kushilevitz, E., Nisan, N., 1997. *Communication Complexity*. Cambridge University Press.
- Li, H., Rosen, S., 1998. Unraveling in matching markets. *Am. Econ. Rev.* 88 (3), 371–387.
- Li, S., 2017. Obviously strategy-proof mechanisms. *Am. Econ. Rev.* 107 (11), 3257–3287.
- McVitie, D.G., Wilson, L.B., 1971. The stable marriage problem. *Commun. ACM* 14 (7), 486–490.
- Ng, C., Hirschberg, D., 1990. Lower bounds for the stable marriage problem and its variants. *SIAM J. Comput.* 19 (1), 71–77.
- Niederle, M., Roth, A.E., 2003. Unraveling reduces mobility in a labor market: gastroenterology with and without a centralized match. *J. Polit. Econ.* 111 (6), 1342–1352.
- Ostrovsky, R., Rosenbaum, W., 2015. Fast distributed almost stable matchings. In: *Proceedings of the 34th ACM Symposium on Principles of Distributed Computing (PODC)*, pp. 101–108.
- Picard, J.-C., 1976. Maximal closure of a graph and applications to combinatorial problems. *Manag. Sci.* 22 (11), 1268–1272.
- Razborov, A.A., 1992. On the distributional complexity of disjointness. *Theor. Comput. Sci.* 106.
- Roth, A.E., 1984. The evolution of the labor market for medical interns and residents: a case study in game theory. *J. Polit. Econ.* 92 (6), 991–1016.
- Roth, A.E., 1986. On the allocation of residents to rural hospitals: a general property of two-sided matching markets. *Econometrica* 54 (4), 425–427.
- Roth, A.E., 1991. A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the United Kingdom. *Am. Econ. Rev.* 81 (3), 415–440.
- Roth, A.E., Xing, X., 1994. Jumping the gun: imperfections and institutions related to the timing of market transactions. *Am. Econ. Rev.* 84 (4), 992–1044.
- Segal, I., 2007. The communication requirements of social choice rules and supporting budget sets. *J. Econ. Theory* 136, 341–378.
- Ünver, M.U., 2005. On the survival of some unstable two-sided matching mechanisms. *Int. J. Game Theory* 33 (2), 239–254.
- Wilson, L.B., 1972. An analysis of the stable marriage assignment algorithm. *BIT Numer. Math.* 12 (4), 569–575.
- Yao, A.C.-C., 1979. Some complexity questions related to distributive computing (preliminary report). In: *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 209–213.