Fair Stable Matching Meets Correlated Preferences

Angelina Brilliantova
Rochester Institute of Technology
Rochester, NY, USA
lb9849@rit.edu

Hadi Hosseini
Pennsylvania State University
University Park, PA, USA
hadi@psu.edu

ABSTRACT

The stable matching problem sets the economic foundation of several practical applications ranging from school choice and medical residency to ridesharing and refugee placement. It is concerned with finding a matching between two disjoint sets of agents wherein no pair of agents prefer each other to their matched partners. The *Deferred Acceptance* (DA) algorithm is an elegant procedure that guarantees a stable matching for any input; however, its outcome may be unfair as it always favors one side by returning a matching that is optimal for one side (say men) and pessimal for the other side (say women). A desirable fairness notion is minimizing the *sex-equality cost*, i.e. the difference between the total rankings of both sides. Computing such stable matchings is a strongly NP-hard problem, which raises the question of what tractable algorithms to adopt in practice. We conduct a series of empirical evaluations on the properties of sex-equal stable matchings when preferences of agents on both sides are correlated. Our empirical results suggest that under correlated preferences, the DA algorithm returns stable matchings with low sex-equality cost, which further confirms its broad use in many practical applications.

KEYWORDS

Stable Matching; Fairness; Correlated Preferences

ACM Reference Format:
Angelina Brilliantova and Hadi Hosseini. 2022. Fair Stable Matching Meets Correlated Preferences. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 9 pages.

1 INTRODUCTION

Matching theory sets the economic foundation for achieving stable allocations through market design. It has shaped the cornerstone of many practical applications ranging from school choice [1, 2] and medical residency [54] to refugee placement [16, 34] and ridesharing [8, 24]. In its essence, the *stable matching* problem deals with finding a matching between two disjoint sets of agents (colloquially men and women) according to their preferences. The primary objective is to achieve stability, that is, finding a matching between the two sides wherein no pair of agents prefer each other to their matched partners.

Over the past few decades, numerous theoretical breakthroughs and developments were pioneered to study mathematical and axiomatic properties of the stable matching problem [17, 21, 53, 59] as well as its computational and algorithmic aspects [26, 36, 40, 57, 62]. However as problems become increasingly more challenging, there has been a need for moving from theory about simple markets to more complex settings that account for subtle, but vital, differences in constraints in preference structures or other factors that may pose computational or axiomatic challenges. Hence, as Roth and Peranson [54] suggested *"as game theory moves from simple conceptual problems to complex design problems, we will need to make more general use of this interaction among theory, computational investigation of market data, and theoretical computation, and that this in turn will produce new problems and directions for traditional theory"*. This doctrine motivated a large body of work in taking empirical or statistical approaches in exploring markets through studying (real or synthetic) data sets or analyzing statistical distributions [6, 11, 39, 61]. In this vein, we investigate the fairness of stable matchings through empirical simulations to paint a thorough picture of the structure of fair stable solutions in matching markets.

A motivating insight. The Deferred Acceptance algorithm (DA)—due to Gale and Shapley [20]—provides an elegant solution to the stable matching problem wherein agents from one side (say men) make proposals to the agents from the other side (say women). Each woman tentatively accepts her favorite proposal and rejects the rest. Despite the popularity and success of the DA algorithm in many real-world matching markets [53, 54], it always favors the proposing side to the receiving side, that is, the DA algorithm always returns a stable matching that is *men-optimal* [20] but *women-pessimal* [48]. This unequal treatment of the sides raises critical questions about the *fairness* of the DA algorithm, which may result in (extremely) unequal welfare between both sides.

One of the most prominent and well-studied fairness notions—proposed by Gusfield and Irving [23]—is *sex-equality* that aims at finding a matching that equalizes the *welfare* of both sides by minimizing the difference between the total rankings of men and women in a stable matching. While sex-equal stable matchings always exist, computing one has shown to be strongly NP-hard [36]. For any instance of the matching problem, there may be an exponential number of stable solutions [31] and these stable matchings form a *distributive lattice*. Thus, one may hope to find a "fair" stable matching that equalizes the welfare of both sides.

On a closer scrutiny, however, we notice that a sex-equal stable matching is highly correlated with the structure of preference lists of each side of the market. This observation suggests that even though the number of stable matchings grows exponentially in general [31], surprisingly a sex-equal solution lies at the extreme points of the stable lattice in certain settings.

Example 1.1. *Consider the following instance with five men and five women and the following preference lists. The men-optimal and women-optimal matchings are marked by † and ∗ respectively. The stable lattice of this preference profile contains five matchings as illustrated in Figure 1. A sex-equal stable matching assigns a total*

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

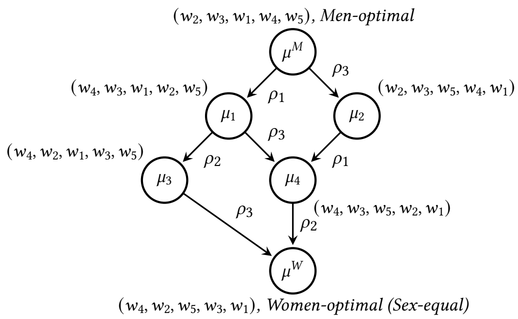

Figure 1: The stable lattice for the profile described in Example 1.1 with six stable solutions. Matchings are denoted as lists of women matched to men according to their indices, rotations are indicated by ρ_x on arcs.

ranking of 12 to women and 13 to men, minimizing the welfare difference between the two sides. Notice that the sex-equal solution for this profile lies precisely on the extreme point of the stable lattice and is equivalent to the women-optimal solution.[1]

m_1 :	$w_2^\dagger$	w_4^*	w_5	w_1	w_3	w_1 :	m_4	m_2	m_1	m_5^*	$m_3^\dagger$
m_2 :	$w_3^\dagger$	w_2^*	w_4	w_1	w_5	w_2 :	m_2^*	m_4	$m_1^\dagger$	m_5	m_3
m_3 :	$w_1^\dagger$	w_5^*	w_4	w_3	w_2	w_3 :	m_4^*	$m_2^\dagger$	m_1	m_3	m_5
m_4 :	$w_4^\dagger$	w_2	w_3^*	w_1	w_5	w_4 :	m_2	m_1^*	$m_4^\dagger$	m_5	m_3
m_5 :	w_2	w_3	$w_5^\dagger$	w_1^*	w_4	w_5 :	m_1	m_4	m_2	m_3^*	$m_5^\dagger$

Motivated by this observation, we study stable matching problems when both sides of the market have correlated preference lists. We focus on correlated random preference lists sampled from *Mallows distribution models* [45]. The Mallows model is the cornerstone of a variety of ranking problems in machine learning [7, 44] and social choice [30], and has been shown to correctly capture the preference rankings of individuals in many practical applications [13, 58]. Under the Mallows models, preferences are correlated through a reference ranking, where the probability of a preference list to be sampled is inversely proportional to its Kendall-tau distance [37] (number of swaps between the two lists) from the reference ranking. Therefore, focusing on settings with correlated preference lists on both sides of the market we ask the following questions:

> *How does the correlation between preference lists of two sides impact the fairness of stable matchings? What algorithmic solutions should we adopt in practice when preferences are correlated according to Mallows models?*

1.1 Our Results

Focusing on markets with correlated preferences on both sides, we empirically investigate the sex-equality of stable matchings through a series of extensive experiments on synthetic data:

- **Stable lattice**: We first focus on the size of the stable lattice when preferences of both sides are drawn from Mallows distributions. We show that the size of the stable lattice is heavily dependent on the relationship between the correlation intensity in preferences of two sides (Section 3). In particular, we show that when correlations between the two sides of the market are *symmetric*, that is, preferences are distributed according to the same dispersion parameter, the stable lattice grows rapidly–as it was theoretically proved by Levy [43]. However, when correlations are *asymmetric*, i.e. distributed according to different dispersion parameters, the size of the stable lattice sharply decreases.
- **Asymmetric correlations**: When preferences of the two sides are sampled from different Mallows distributions, we show that in overwhelming majority of cases, a sex-equal stable matching is located at the extreme points of the stable lattice. In Section 4.1 we discuss how this key observation immediately results in a polynomial-time algorithm for finding a sex-equal matching by computing a men-optimal or women-optimal stable matching through the DA algorithm [20].
- **Symmetric correlations**: When preferences are sampled from the same Mallows distribution, even though the size of the stable lattice may be exponential, the cost of a sex-equal matching is considerably close to the cost of the DA outcome. As the difference between the welfare of men and women is sufficiently small, the importance of which group proposes (men or women) becomes negligible (Section 4.2).
- **Comparing the performance of algorithms**: In Section 5 we conduct a series of empirical comparisons between well-studied heuristic and procedurally fair algorithms. We show that with correlated preferences, the DA algorithm—with a careful selection of the proposing side—performs as good as the best known local search algorithm with respect to the sex-equality cost even on large instances. This result further justifies the use of the DA algorithm which is more computationally efficient compared to other methods.

1.2 Related Work

Stability is a key condition for the success and longevity of two-sided markets [53]. The seminal work by Gale and Shapley showed that a stable matching always exists and can be found in polynomial time using the Deferred Acceptance algorithm (DA) [20]. The important property of DA is its inherent asymmetry: favouring one side at the cost of another [23]. This bias should be taken into account as many central clearinghouses use DA as the basis for their decisions, including the National Resident Matching Program (NRMP) [54] and the New York City school assignment [1].

Several papers investigate the difference between the expected rank of partners in an optimal and a pessimal stable matching (aka *the welfare gap*) [5, 6, 51, 52]. In one-to-one balanced markets (i.e. with an equal number of men and women) and when preferences drawn uniformly at random, the expected total rank of agents asymptotically approaches $n \log n$ in their optimal matching and $n^2/\log n$ in their pessimal matching [51]. The same order of scores is preserved in markets with constant tier-based preferences, in which agents rank partners proportionally to their real-numbered popularity scores [5]. Large real-world markets usually admit an exceedingly small number of stable matchings. For example, for

[1] We refer the readers to the full version of the paper [12] for a detailed discussion of lattices and rotations.

the NRMP hospital-resident matching and for Boston Public School student admission there were only a couple of stable matchings each year [50, 54]. It was shown theoretically, that a lattice becomes essentially a singleton in large one-to-one markets, wherein one side draws shorter fixed-sized preference lists from an arbitrary distribution [28]. Similar results were shown for many-to-one settings: balanced [41] and unbalanced [6]. At the same time, large stable lattices were theoretically proven in markets with correlated preferences induced by the Watts–Strogatz "small world" model (even when unbalanced) [52], and the Mallows model [43]. Empirically, large lattices are found in settings of matching-with-contracts [25].

Fairness and stability. The potential exponential size of stable solutions [31] alongside the inherent bias of DA, has motivated the discussion of fairness in stable matching markets. Various algorithms were proposed to ensure fairness towards individuals [18, 19], socio-economic groups [3, 27], regions [35], and sides (e.g. hospitals vs. residents) [47, 60]. Below, we will briefly discuss a variety of fairness notions that have received attention in the literature of stable matchings.

The *egalitarian* stable matching, is a stable matching that minimizes the sum of rankings of all partners for men and women, hence maximizes the total welfare of all agents [32]. The *minimum regret* stable matching minimizes the highest rank of agents, therefore minimizing the rank of a partner for the most unsatisfied agent. Egalitarian and minimum regret stable matchings can be found in polynominal time [18, 32], but do no ensure cross-sided fairness: they both can substantially favor one of the sides, by optimizing the welfare on the level of the whole system (egalitarian), and individual agents (minimum regret). In the *median stable matching*, each agent is matched to its middle favorite partner across all partners from stable matchings; the resulting matching is a median element of a stable lattice. Finding a median stable matching is NP-hard in general, but it could be done in polynomial time in some restricted families of the rotation poset [14]. The median stable matching might not satisfy cross-sided fairness if one side prefers their middle stable partners substantially more than the other side. A *procedurally fair* algorithm aims to provide agents an equal probability to affect the resultant stable matching by issuing proposals [61], reducing the set of stable matchings [4], or satisfying blocking pairs [55]. Procedural fairness does not necessarily result in cross-sided fair matchings, and thus, one side may receive much higher welfare compare to the other side [38].

A *sex-equal stable matching* minimizes the gap between the sum of partners' ranks of men and women. The sex-equal stable matching problem is an NP-hard problem and fixed parameter tractable with respect to the treewidth of the Hasse diagram of the rotation poset [22] and k parameter in the k-range model [15]. To tackle the sex-equal stable matching problem various heuristics have been proposed: for example, performing local search series on a stable lattice (iBILS) [64], and transforming a stable matching using genetic algorithm [49]. While some of these approximate algorithms (e.g. iBILS) perform well in experiments, they bear no theoretical bounds on the quality of an outcome.

Correlated and random preferences. Several recent works in stable matching have focused on investigating a variety of natural structures of profiles, from profiles generated uniformly at random [6, 51] to random profiles with soft and hard constraints on rankings [5, 39, 61]. The Mallows distribution model drew the attention of researchers as it was shown to realistically capture the preference of agents in several applications involving individual decision-makers [58]. In the Mallows model, preferences are correlated through a reference ranking, representing the objective order of agents' attractiveness. In contrast to the tiered and random utility models, the Mallows model generates a variety of correlated preferences using only a few parameters: two dispersion parameters and two reference rankings; one of each for each side of the market.

2 PRELIMINARIES

We start by providing a formal representation of the model and define the necessary properties.

Problem setup. An instance of the *stable matching problem* is specified by the tuple $I = \langle M, W, \succ\rangle$, where M is a set of n men, W is a set of n women, and $\succ$ is a *preference profile* which consists of the preference lists of all men and women. The preference list of any man $m \in M$, denoted by $\succ_m$, is a strict total order over all women in W (for any woman $w \in W$, the list $\succ_w$ is defined analogously). The rank of woman w in man m's preference list, $\succ_m$, is denoted by $r(w, m)$. For instance, given $\succ_m = (w_2, w_1, w_3)$, we say w_2 is ranked first in m's preference list, i.e., $r(w_2, m) = 1$. Similarly, $r(m, w)$ denotes the rank of m in $\succ_w$.

Stable matchings. A perfect *matching* is a mapping $\mu : M \cup W \to M \cup W$ such that $\mu(m) \in W$ for all $m \in M$, $\mu(w) \in M$ for all $w \in W$, and $\mu(m) = w$ if and only if $\mu(w) = m$. Given a matching μ, a man-woman pair (m, w) is called a **blocking pair**, with respect to the preference profile $\succ$, if they prefer each other to their assigned partners under μ, i.e., $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching is **stable** if it contains no blocking pairs.

Given a preference profile $\succ$, the set of all corresponding stable matchings, $S_\succ$, forms a *distributive lattice*. The maximum and minimum points of a stable lattice correspond to the men-optimal (μ^M) and the women-optimal (μ^W) matchings respectively [23], i.e. a matching where all men (respectively women) receive their best stable partners. A men-optimal matching is simultaneously women-pessimal: it matches all woman to their worst stable partners [48]. Figure 1 illustrates a lattice consisting of six stable matchings. The expected size of a stable lattice (correspondingly the size of $S_\succ$) is asymptotic to $e^{-1} \ln n$ when preferences are drawn uniformly at random [51] and may grow exponentially with the number of agents [31].

Welfare measure. Let $S_M(\mu)$ denote the sum of ranks of men's partners in matching μ, that is, $S_M(\mu) = \sum_{m \in M} r(\mu(m), m)$. Similarly for women, we let $S_W(\mu)$ be the sum of ranks of women's partners given matching μ, i.e. $S_W(\mu) = \sum_{w \in W} r(\mu(w), w)$. Note that the smaller values indicate higher social welfare. Thus, a men-optimal matching μ^M is the one that minimizes the sum of rankings for men, that is, $\mu^M = \arg\min_{\mu \in S_\succ} S_M(\mu)$. And similarly, for the women-optimal matching we have $\mu^W = \arg\min_{\mu \in S_\succ} S_W(\mu)$. For simplicity, we will refer to the scores of a men-optimal matching $S_M(\mu^M)$, $S_W(\mu^M)$ as S_M**-optimal** and S_W**-pessimal** scores. Similarly, for a women-optimal matching we write S_M**-pessimal** and S_W**-optimal** to indicate $S_M(\mu^W)$ and $S_W(\mu^W)$.

The Deferred Acceptance algorithm. Given a preference profile $\succ$, the Deferred Acceptance (DA) algorithm, proposed by Gale and Shapley [20], consists of rounds of *proposal* and *rejection* phases and proceeds as follows: In each round every man who is currently unmatched proposes to his favorite woman from among those who have not rejected him yet. Each woman tentatively accepts her favorite proposal and rejects the rest. The algorithm terminates in $O(n^2)$ when no further proposals can be made.

Given any profile $\succ$ as input, the DA algorithm is guaranteed to return a stable matching [20]. Moreover, the DA algorithm is simultaneously *men-optimal* [20] and *women-pessimal* [48], i.e. men receive their favorite stable partners among all matchings available to them in $S_\succ$, and women receive their least favorite stable partners. The DA algorithm returns the stable matchings at the extreme points of the stable lattice depending on which set (men or women) are proposers. Thus, the DA algorithm with women proposing is respectively women-optimal and men-pessimal.

2.1 Fair Stable Matchings

Given a matching μ, the *sex-equality cost* of μ is the absolute difference between the total welfare of men and women, i.e. S_M and S_W scores. Formally, the sex-equality cost of a matching μ is defined as

$$c(\mu) = |S_M(\mu) - S_W(\mu))|. \tag{1}$$

Given preference profile $\succ$, a *sex-equal stable matching* is a matching in $S_\succ$ that minimizes the sex-equality cost across the stable lattice, that is,

$$\mu^* \leftarrow \arg\min_{\mu \in S_\succ}(c(\mu)). \tag{2}$$

Intuitively, in a sex-equal stable matching the total rank of men's partners is as close as possible to that of women (subject to the stability condition). Kato [36] showed that for an arbitrary preference profile the problem of finding a sex-equal stable matching is strongly NP-hard. However, it is fixed parameter tractable with respect to the range of the profile – a metric showing the maximum discrepancy between an agent's worst and best rank in the preference profile [15]. Also, it can be found efficiently when agents have a specific two-dimensional single-peaked model [56] or if preferences are identical on one side [33].

2.2 Correlated Preferences

There are several plausible ways to study preference models generated from uniform distributions [6, 51] or correlated preferences [5, 9, 52, 61]. The vast majority of these works focused on profiles wherein the preferences of one side are correlated while the other side is considered uniform. We focus on a more *general* distribution models where *both* sides of the market have correlated preferences through the Mallows model.

The Mallows model. A Mallows distribution is a distance-based probabilistic model for permutations correlated with some common reference [45]. It is parameterized by a reference ranking, $\hat{\pi}$, and a dispersion parameter $\phi \in (0, 1]$. Let π be a permutation of a preference list. For any permutation π, the Mallows model specifies a probability as follows:

$$p(\pi|\hat{\pi}, \phi) = \frac{1}{Z}\phi^{\tau(\pi,\hat{\pi})}$$

where $\tau(\pi, \hat{\pi})$ is a Kendall-tau distance (the number of pairwise inversions) between π and $\hat{\pi}$ and Z is a normalization constant with $Z = 1(1+\phi)(1+\phi+\phi^2)\dots(1+\phi+\dots+\phi^{n-1})$.

Note that the dispersion parameter ϕ indicates the 'intensity' of correlation between the sampled preferences. When $\phi = 0$, the correlation is maximal, that is, the distribution mass is entirely on the reference ranking and all preference lists are identical. When $\phi = 1$, the correlation is minimal, the probability mass is distributed evenly between all possible permutations and the Mallows model is equivalent to the **Uniform distribution model** (also known as *Impartial Culture* [10] in the computational social choice literature).

For each set involved in the stable matching problem, we consider an independent probabilistic preference model: one Mallows model for the set of men parameterized by $\hat{\pi}_m$ and ϕ_m and another preference model for women specified by $\hat{\pi}_w$ and ϕ_w. In this way, the preferences of men (similarly women) are globally *correlated* with $\hat{\pi}_m$ ($\hat{\pi}_w$).[2]

Symmetric and asymmetric models. We let $\phi_\Delta = |\phi_m - \phi_w|$ to represent the *correlation disparity* in preferences of men and women sampled for this particular instance. Simply put, correlation disparity reflects how similar the probabilistic preferences of men are *in comparison* to the similarity of women's preferences. We call stable matching instances simulated from the Mallows model with zero correlation disparity, $\phi_\Delta = 0$, *symmetric correlation markets*, and *asymmetric correlation markets* when simulated with $\phi_\Delta \neq 0$.

3 STABLE LATTICE UNDER THE MALLOWS MODEL

In this section, we show empirically that in one-to-one markets with correlated preferences induced by Mallows models the size of the stable lattice depends on the correlation disparity. The empirical investigations give insights on how searching for a sex-equal matching can computationally vary with the size of a stable lattice in a given market.[3]

3.1 Setup and Preference Sampling

We generate 1,000 instances of a stable matching problem for each $n \in [10, 150]$ and ϕ_m, ϕ_w in $[0.1, 1.0]$. When generating a preference profile under probabilistic models, we draw a preference list for each agent *i.i.d* from the Mallows distribution using the sampler from PrefLib library [46].

In all sampled stable matching instances we built the stable lattice, and measured S_M, S_W welfare scores. For each experimental setting, we estimate a target statistic (the median for numerical variables like size and mean for binary variables) over 1,000 instances along with the confidence intervals. Confidence intervals were obtained using *bootstrap sampling with replacement* with sample size 1,000 with 100 repeats. We report the results only for $n = 150$, as for the smaller n the results are qualitatively the same. We occasionally considered larger instances with $n = 300$, when comparing the size of the lattice of the Uniform and the Mallows ($\phi_m = \phi_w = 0.5$) models.[4]

[2] Throughout the paper, we assume $\hat{\pi}_m = \hat{\pi}_w$, because one can simply relabel the preferences on one side to create any arbitrary distance.
[3] The source code is available at https://github.com/Restel/mallows-smp
[4] We use the same instance generation and setup in all remaining sections.

	n	Type	Median $\mathcal{L}$	CI($\mathcal{L}$)	max($\mathcal{L}$)
1	150	Mallows	16.00	(2;256)	1024
3	150	Uniform	82.00	(32;231)	626
4	300	Mallows	144.00	(8;7008)	147456
5	300	Uniform	209.00	(91;567)	1262

Table 1: The size of the stable lattice ($\mathcal{L}$) under the Uniform, and the Mallows distributions with $\phi_m = \phi_w = 0.5$ models. $CI(\mathcal{L})$ denotes the 95% confidence interval.

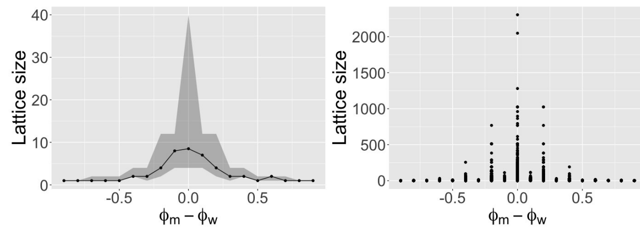

Figure 2: The size of the stable lattice with respect to $\phi_m - \phi_w$ in the Mallows model for $n = 150$: median values (left), maximum values (right). Dotted lines denote the medians, shadowed gray area denotes 1st and 3rd quartiles. Y-axes of left and right figures have different scales.

3.2 The Size of the Stable Lattice

The size of the stable lattice depends on the correlation disparity. Our empirical investigation illustrates an intriguing observation about the size of the stable lattice: Under the Mallows model the size of the stable lattice is affected by the correlation intensity in men's and women's preferences (measured by the dispersion parameters ϕ_m, ϕ_w) as well as the discrepancy between them, i.e. the correlation disparity ϕ_Δ. The size of the lattice is maximum when the preferences of men and women are sampled from a distribution with identical dispersion parameters, $\phi_\Delta = 0$, and decreases with the rise of the correlation disparity. This relationship is confirmed by median values, 25%-75% percentiles, 95% confidence intervals, maximum values and holds for all combinations of dispersion parameters (Figure 2). These findings are aligned with theoretical results indicating that when the preferences are sampled from the same distribution (parametrized by identical dispersion parameters, $\phi_m = \phi_w$), the size of the stable lattice grows exponentially in the number of agents [43], making the use of an exhaustive search algorithm impractical in these settings.

More importantly, they reveal that the primary predictor of the lattice size is the correlation disparity between the two sides, and the intensity of the correlations (dispersion parameters) have less significant impact on the lattice size.[5]

Mallows preferences induce large lattices. It is thought that introducing correlation in preferences reduces the stable lattice size compared to the random uncorrelated case [6, 42, 54]. In their empirical work, Roth and Peranson [54] conjectured that having a common objective ranking for the agents' preferences is one of the main reason behind the observed small lattice sizes. Our empirical results suggest that while in general the median size of the lattice was indeed smaller in the Mallows compared to the Uniform distribution, some ϕ parameters have much larger maximum values and confidence intervals.[6] For example, in case of asymmetric correlation with $\phi_m = 0.5$, $\phi_w = 0.3$, $n = 150$, the maximum size of the lattice was larger than that of the Uniform model. Similarly, for symmetric case with $\phi_m = 0.5$, $\phi_w = 0.5$, and $n = 300$, 95% confidence interval of a stable lattice size spans from 8 to 7008 matchings compared to 91–567 interval under the Uniform distribution (see Table 1).

Our results suggest that symmetrically correlated markets have extraordinarily large lattices compared to the Uniform case (147456 vs. 1262), aligning well with the results of Levy [43] on asymptotic exponential lattices in the Mallow model. Note, that such large instances appeared for such small number of agents as 300, while real-world markets often have many more agents [50, 54].

In the past, correlation in preferences of agents was considered a sole factor contributing to small lattice sizes [6, 54]. Our experiments imply that in case of the correlation induced by the Mallows model, this effect depends not on the presence of correlation itself, but rather on its disparity between the two sides i.e. *how strongly the preferences of one set are correlated compared to the preferences of the other set.* Symmetric correlation markets seem to be the worst case scenario, and the lattice tends to grow smaller when the sets are having preferences correlated dissimilarly. This raises the question of whether real-world markets with observed small lattices suffer from such correlation disparity in preference profiles or it should be attributed to other factors.

4 SEX-EQUAL STABLE MATCHINGS

In this section, we discuss how the position of a sex-equal matching within a stable lattice is determined by the relationship between men and women welfare scores, which enables us to find a sex-equal matching using the DA algorithm.

The relationship between S_M and S_W of men- and women-optimal stable matchings dictates the location of a sex-equal solution within the lattice. In particular, it determines whether or not it lies on the extreme points of the lattice. Based on this relationship, we identify two "easy" cases, in which DA guarantees to return a sex-equal stable matching, and a "hard" case, in which it does not.

Lemma 1 (Kato [36]). *A sex-equal stable matching is:*
1) the men-optimal stable matching, μ^M, if $S_M(\mu^M) \geq S_W(\mu^M)$
2) the women-optimal matching, μ^W, if $S_M(\mu^W) \leq S_W(\mu^W)$
3) the stable matching μ_Z such that $c(\mu_Z) = \arg\min_{\mu \in S_>}(c(\mu))$, otherwise.

Corollary 1. *A sex-equal stable matching can be computed in polynomial time when either $S_M(\mu^M) \geq S_W(\mu^M)$ or $S_M(\mu^W) \leq S_W(\mu^W)$.*

Proof. Find a men-optimal matching μ^M by running a men-proposing DA algorithm. Compute $S_M(\mu^M)$ and $S_W(\mu^M)$ and if $S_M(\mu^M) \geq S_W(\mu^M)$, than by Lemma 1, Case 1 the men-optimal matching is a sex-equal stable matching. Analogously, find a women-optimal matching μ^W by running a women-proposing DA algorithm. Compute $S_M(\mu^W)$ and $S_W(\mu^W)$ and if $S_M(\mu^W) \leq S_W(\mu^W)$, than

[5] Similar results hold for rotation posets as described in the full version [12].
[6] The details are provided in [12].

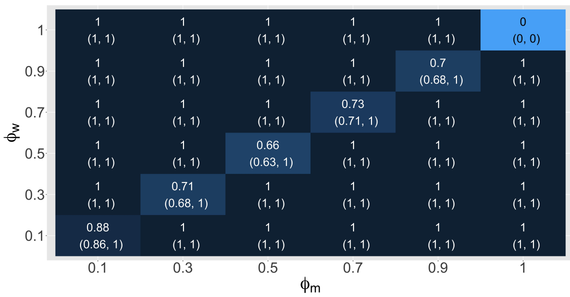

Figure 3: The fraction of instances where a sex-equal solution is an outcome of DA algorithm for $n = 150$: men proposing if $\phi_m < \phi_w$, women-proposing if $\phi_m > \phi_w$, any side proposing if $\phi_m = \phi_w$. Confidence intervals are given in parentheses.

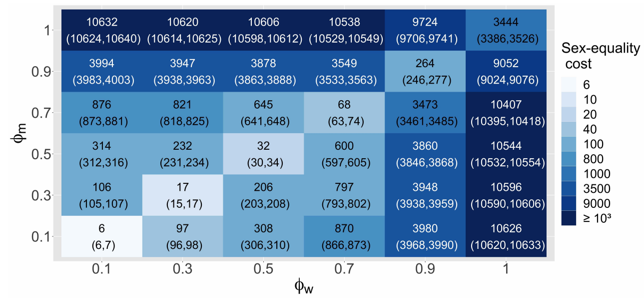

Figure 4: The median sex-equality cost of a DA solution with respect to ϕ_m, ϕ_w, for $n = 150$. Confidence intervals are given in parentheses.

by Lemma 1, Case 2 the women-optimal matching is a sex-equal stable solution. The running time of DA and calculating S_M, S_W scores is $O(n^2)$. Thus, a sex-equal stable matching can be found in polynomial time when $S_M(\mu^M) \geq S_W(\mu^M)$ or $S_M(\mu^W) \leq S_W(\mu^W)$. □

In the first and the second cases of Lemma 1, the sign of the difference between the scores $S_M(\mu) - S_W(\mu)$ is preserved across the lattice, in other words $S_M(\mu)$ is either always greater or equal than $S_W(\mu)$ (Case 1), or smaller or equal than $S_W(\mu)$ (Case 2). The third case corresponds to the instances, in which the sign of $S_M(\mu) - S_W(\mu)$ changes across the lattice, and the location of the sex-equal matching within the lattice is not known a priori.

4.1 Asymmetric Markets ($\phi_m \neq \phi_w$)

When preferences of men and women have different levels of correlation intensity with the reference rankings, a sex-equal stable matching is often found on an extreme point of the stable lattice: men-optimal in case men have higher correlation and women-optimal otherwise. Empirically, we observe this behavior in all instances, including the largest stable lattices with up to 1024 matchings (Figure 3). We found that agents from the side with a smaller dispersion parameter ϕ of the Mallows model, are less satisfied with their optimal matching than their partners with their pessimal one (and the market belongs to Case 1 or Case 2 scenario from Lemma 1). This fact might imply that, in asymmetric correlation markets, "hard" cases rarely occur and the DA algorithm can be considered a fair alternative to the exhaustive search, especially since such markets can induce large lattices as shown in Section 3.

Experimentally, in every stable matching instance belonging to an asymmetric correlation market with $\phi_m < \phi_w$, S_M-optimal was larger than S_W-pessimal, and therefore a men-optimal matching was sex-equal (Lemma 1, Case 1). Similarly, whenever $\phi_m > \phi_w$, S_W-optimal was larger than S_M-pessimal, rendering the women-optimal matching sex-equal (Lemma 1, Case 2). This is well illustrated by the distribution of S_M and S_W scores of men- and women-optimal matchings for various combinations of ϕ_m, ϕ_w. Figure 5 illustrates this relationship when men sample their preferences with $\phi_m = 0.9$, and women have their dispersion parameter in range $[0.1, 1.0]$ (the graphs for other values of ϕ_m are analogous).

In the part of the graph, where women's preferences have a lower dispersion parameter $\phi_w < \phi_m$, the distribution of S_W-scores of a women-optimal matching lies strictly above the distribution of its S_M-scores with no overlap.[7]

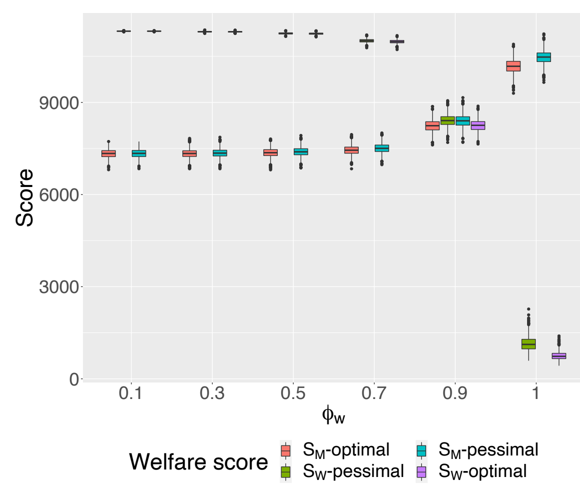

Figure 5: Illustrating S_M, S_W scores under the Mallows model with $\phi_m = 0.9$ and varying ϕ_w based on 1000 samples.

Implications on the DA algorithm. Interestingly, increasing the variability of women's preferences, results in 1) higher welfare for women in optimal and pessimal matchings and 2) a higher number of proposals from men. The total number of proposals in the DA algorithm corresponds to the optimal score of the proposing side [51]. In Figure 5 we can see a noticeable increase in the S_M-scores with the increase of ϕ_w as well as a decrease in S_W scores (same pattern occurs for other values of ϕ_m). Imagine that the proposing side samples its preferences from the model with a smaller dispersion parameter, and hence has less variable preferences and

[7] A more fine-grained analysis with $\phi_w \in \{0.9, 0.91, 0.92, \ldots, 0.99\}$ confirms the same finding (see the full paper for details [12]).

demonstrates a higher competition for partners (than the opposite side). Intuitively, during the course of DA those agents would propose to the same set of partners over and over again and their potential partners get to choose from a diverse set of proposals, which improves the rank of their matches. This makes the total rank of proposers' partners greater than that of the agents from the accepting side. This fact distinguishes the Mallows model from the Uniform model in which the proposing side gets a substantially smaller score than the accepting side [51].

4.2 Symmetric Markets ($\phi_m = \phi_w$)

When the preferences of men and women have the same intensity of correlation, reflected in zero correlation disparity, the lattice size can quickly become very large. The size of the stable lattice is expected to be asymptotically exponential [43], which is confirmed in our experiments: even for small number of agents as $n = 300$, there were significantly large (compared to the Uniform model) lattices with up to 147456 matchings (Table 1). Despite the large size of the stable lattice, the cost of a sex-equal solution is considerably close to the cost of DA outcomes (Figure 4). In this case, the DA algorithm can be considered cost-effective as the exhaustive search takes significantly more time for large lattices and gives only a moderate gain in the sex-equality cost.

We discussed in Section 4.1, that under the Mallows model, S_M-optimal and S_W-pessimal scores (similarly S_W-optimal, S_M-pessimal) are close to each other and their distributions overlap significantly. This can explain the small sex-equality cost of DA outcomes (see Figure 5). It could indicate, that under this scenario, the expected score of the proposing side under DA is similar to the expected score of the receivers. This finding differs sharply from the Uniform model wherein the expected pessimal score is substantially greater than the optimal one [5, 51]. As the difference between the welfare of men and women is sufficiently small, the importance of which group proposes (men or women) in DA becomes negligible. Hence, not only can the DA algorithm be adopted to ensure sex-equality, but it may also provide a natural framework for achieving *procedural fairness* as it has recently been studied [60, 61].

4.3 Implications on manipulation strategies

In our experiments, the Mallows model with symmetric and asymmetric correlations results in a considerably small difference between the pessimal and the optimal scores for one side (*the welfare gap*), for example, between S_W pessimal and S_W-optimal (Figure 5). This difference represents how much on average an agent can improve by shifting from her worst stable partner to her best one, and therefore has implications on manipulation incentives [6, 52]. Since the welfare gap is small for both men and women, in symmetric and asymmetric correlation settings, the accepting side has a limited scope/incentive for manipulations in practice. This observation suggests that one-to-one markets with Mallows preferences might be less susceptible to manipulation compared to the Uniform model, for which a larger welfare gap, and subsequently, higher incentive for manipulation has been established [51] and shown empirically [26, 59].

5 COMPARING WITH OTHER ALGORITHMS

We compare the DA algorithm with several state-of-the art algorithms (as implemented in [61]) such as multiple variants of the procedurally fair algorithms, and the best known heuristic algorithm when preferences are drawn from the Mallows model.[8] We use an *exhaustive search* algorithm as our baseline to find sex-equal stable matchings [60]. In particular, we test several procedures allowing both sides to issue and receive proposals such as *Late Discontent Dispension (LDS)*, *Early Discontent Dispension*, and *Powerbalance* [61]. We also measure the performance of an enhanced *bidirectional local search iBILS*, which is one of the best-performing existing heuristic for finding sex-equal stable matchings.

5.1 Experimental Setup

We evaluate the sex-equality cost and the run-time of these algorithms for the number of agents varying from 20 to 1000. For the symmetric correlation case, we select $\phi_m = \phi_w = 0.5$ as these parameters results in the largest lattice size among all combinations of ϕ_m and ϕ_w. We sample 1000 instances of the stable matching for each experimental setting.

The *exhaustive search* algorithm uses the *break-marriage* operation of McVitie and Wilson [48]: breaking a single man's assignment initiates a chain of proposals and rejections. It terminates when either all women reject the last man, or his proposal is accepted by a single woman. The latter case results in a new stable matching. All the elements of the stable lattice can be found by applying this operation recursively from the men-optimal solution [48].[9]

The *enhanced bidirectional local search*, iBILS, uses rotations to generate the successors of a stable matching to search over the stable lattice starting from its extreme points [60, 63].

5.2 Fairness and Running Time

Symmetric correlation. For symmetric correlation markets we used a smaller number of agents (up to 150), to avoid instances with extremely large lattices, for which exhaustive search would not terminate in a reasonable time (Section 4.2). Under symmetric correlations, stable lattice can grow exponentially [43]. In our experiments, all algorithms performed similarly on average with respect to the sex-equality cost (Figure 6). Among them, the iBILS algorithm performed better, and in most instances, found the sex-equal stable matching that was computed by the exhaustive search algorithm. These results and the good performance of other algorithms are justified by the fact that under the Mallows model, S_M and S_W scores of all stable matchings are close to one another (i.e. the welfare gap between the two sets is small). Therefore, many stable matchings have costs close to that of the the sex-equal solution.

While the DA and procedurally fair algorithms give higher sex-equality costs (on average twice as much as the exhaustive search and the iBILS algorithms), they have much smaller running times. For larger instances, this discrepancy between the running times will become even more pronounced. The exhaustive search has exponential complexity in the worst case. The running time of the

[8]The comparison of algorithms for the Uniform model can be found in the full paper [12].

[9]An example of the break-marriage operation is provided in the full paper [12].

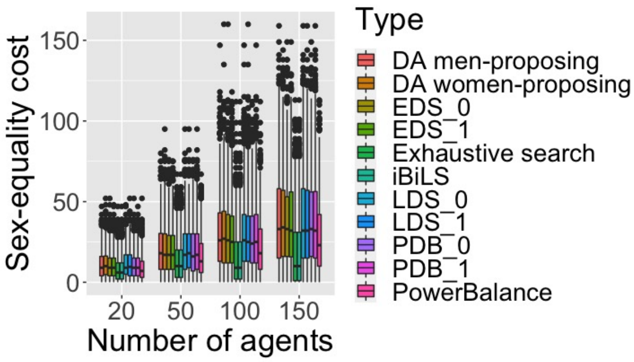

Figure 6: Mallows model (symmetric): comparing the sex-equality cost of various matching algorithms when $\phi_m = \phi_w = 0.5$. Subscript 0 indicates that men start proposing first, and 1 indicates women propose first.

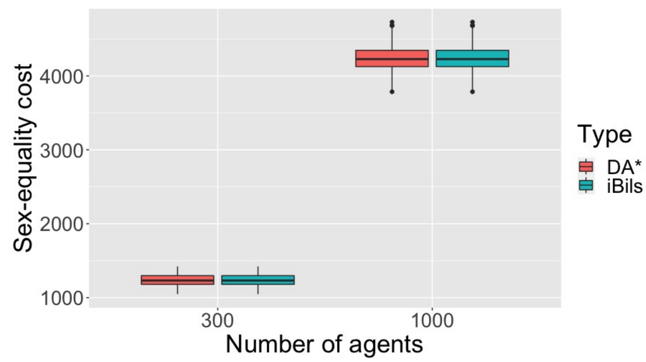

Figure 7: Mallows model (asymmetric): the sex-equality cost of iBILS and a pre-processed Deferred Acceptance (DA*) for an asymmetric correlation market with $\phi_m = 0.5$, $\phi_w = 0.7$.

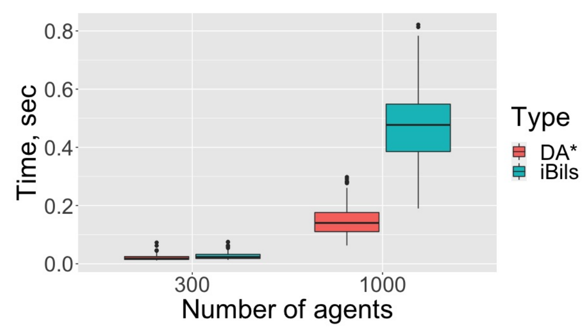

Figure 8: Mallows model (asymmetric): the running time of iBILS and a pre-processed Deferred Acceptance (DA*) for an asymmetric correlation market with $\phi_m = 0.5$, $\phi_w = 0.7$ (other combinations of parameters give analogous results).

iBILS algorithm depends on the stable lattice size and its configuration: it is $O(dkn^2)$, where d is the maximum search depth and k is the maximum width of the lattice [60, 64]. The DA algorithm and procedurally fair algorithms run in polynomial time.

Asymmetric correlation. In asymmetric correlation markets, we focused on comparing the performance of iBILS and DA with additional pre-processing step (aka DA*). Given the findings in Section 4.1, we modify the DA algorithm by deciding which side should propose according to the dispersion parameters: the side with smaller dispersion parameter proposes, i.e. when $\phi_m < \phi_w$ men propose, otherwise women propose. In real-world markets, dispersion parameters can be inferred from preference profiles in polynomial time when the reference ranking is known [29].

Figure 7 presents the sex-equality cost of DA* compared to the iBILS algorithm for large markets with 300 and 1000 agents. In almost all instances, DA* finds the same stable matching as iBILS. More importantly, as shown in Figure 8, the run-time of the DA* remains in $O(n^2)$ while iBILS requires more steps to estimate a sex-equal solution. Given the performance of iBILS in most cases, our results suggest that DA* is a reasonable cost-effective fair algorithm for matching markets with asymmetric correlations.

6 CONCLUDING REMARKS

For decades, market design has played a crucial role in setting the foundations of decision-making through mechanism design. While the theoretical guarantees sometimes fall short due to incompatibilities between axiomatic requirements or challenges imposed by computational complexity, empirical evaluations can still shed light on practical implications and guide the design of new mechanisms. This in turn could help produce new problems and directions for traditional theory.

In this spirit, our extensive empirical investigations revealed an intriguing relation between the disparity of preferences of both sides and the fairness of stable matchings. We showed that surprisingly, the primary factor affecting the sex-equality cost of the DA algorithm is the difference between dispersion parameters of both sides. These observations suggest that the DA algorithm could still be a prime candidate both in symmetric and asymmetric markets, further justifying its wide use in practice.

From the theoretical perspective, investigating the bounds of fairness in stable matchings under correlated preferences is certainly an interesting future direction. From the practical perspective, a noteworthy, and perhaps more important, direction is to investigate the structure of preferences in various matching settings (e.g. refugee matching, residency matching, and school choice) and develop domain-specific models that can correctly capture the correlations between the preferences with high accuracy. We believe that these models, alongside with theoretical studies, should shape governmental and societal policies—prescribed by social planners—in the adoption of suitable mechanisms in each specific domain.

ACKNOWLEDGMENTS

Hadi Hosseini acknowledges support from NSF IIS grants #2052488 and #2107173. We are grateful to the anonymous referees for their valuable comments and suggestions.

REFERENCES

[1] Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. 2005. The New York City High School Match. *American Economic Review* 95, 2 (2005), 364–367.
[2] Atila Abdulkadiroğlu, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez. 2005. The Boston Public School Match. *American Economic Review* 95, 2 (2005), 368–371.
[3] Atila Abdulkadiroğlu and Tayfun Sönmez. 2003. School choice: A mechanism design approach. *American economic review* 93, 3 (2003), 729–747.
[4] Brian Aldershof, Olivia M Carducci, and David C Lorenc. 1999. Refined inequalities for stable marriage. *Constraints* 4, 3 (1999), 281–292.
[5] Itai Ashlagi, Mark Braverman, Clayton Thomas, and Geng Zhao. 2020. Tiered Random Matching Markets: Rank is Proportional to Popularity. *arXiv preprint arXiv:2009.05124* (2020).
[6] Itai Ashlagi, Yash Kanoria, and Jacob D Leshno. 2017. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy* 125, 1 (2017), 69–98.
[7] Pranjal Awasthi, Avrim Blum, Or Sheffet, and Aravindan Vijayaraghavan. 2014. Learning mixtures of ranking models. In *Advances in Neural Information Processing Systems*. 2609–2617.
[8] Siddhartha Banerjee and Ramesh Johari. 2019. Ride sharing. In *Sharing economy*. Springer, 73–97.
[9] Hedyeh Beyhaghi, Daniela Saban, and Eva Tardos. 2017. Effect of selfish choices in deferred acceptance with short lists. *arXiv preprint arXiv:1701.00849* (2017).
[10] Duncan Black. 1958. The theory of committees and elections. (1958).
[11] James W Boudreau and Vicki Knoblauch. 2010. Marriage matching and intercorrelation of preferences. *Journal of Public Economic Theory* 12, 3 (2010), 587–602.
[12] Angelina Brilliantova and Hadi Hosseini. 2022. Fair Stable Matching Meets Correlated Preferences. arXiv:2201.12484 [cs.GT]
[13] Ioannis Caragiannis, Ariel D Procaccia, and Nisarg Shah. 2013. When do noisy votes reveal the truth?. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*. 143–160.
[14] Christine T Cheng. 2010. Understanding the generalized median stable matchings. *Algorithmica* 58, 1 (2010), 34–51.
[15] Christine T Cheng and Will Rosenbaum. 2021. Stable matchings with restricted preferences: Structure and complexity. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. 319–339.
[16] David Delacretaz, Scott Kominers, and Alexander Teytelboym. 2019. *Matching Mechanisms for Refugee Resettlement*. Technical Report. Human Capital and Economic Opportunity Working Group.
[17] Gabrielle Demange, David Gale, and Marilda Sotomayor. 1987. A further note on the stable matching problem. *Discrete Applied Mathematics* 16, 3 (1987), 217–222.
[18] Tomás Feder. 1992. A new fixed point approach for stable networks and stable marriages. *J. Comput. System Sci.* 45, 2 (1992), 233–284.
[19] Rupert Freeman, Evi Micha, and Nisarg Shah. 2021. Two-sided matching meets fair division. *arXiv preprint arXiv:2107.07404* (2021).
[20] David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.
[21] David Gale and Marilda Sotomayor. 1985. Some Remarks on the Stable Matching Problem. *Discrete Applied Mathematics* 11, 3 (1985), 223–232.
[22] Sushmita Gupta, Saket Saurabh, and Meirav Zehavi. 2017. On treewidth and stable marriage. *arXiv preprint arXiv:1707.05404* (2017).
[23] Dan Gusfield and Robert W Irving. 1989. *The stable marriage problem: structure and algorithms*. MIT press.
[24] Jonathan V Hall, John J Horton, and Daniel T Knoepfle. 2017. Labor market equilibration: Evidence from uber. (2017). Working paper. Revision requested from American Economic Review.
[25] Avinatan Hassidim, Assaf Romm, and Ran I Shorrer. 2018. Need vs. merit: The large core of college admissions markets. *Merit: The Large Core of College Admissions Markets (June 28, 2018)* (2018).
[26] Hadi Hosseini, Fatima Umar, and Rohit Vaish. 2021. Accomplice Manipulation of the Deferred Acceptance Algorithm. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*. 231–237.
[27] Chien-Chung Huang. 2010. Classified stable matching. In *Proceedings of the twenty-first ACM-SIAM symposium on Discrete Algorithms*. SIAM, 1235–1253.
[28] Nicole Immorlica and Mohammad Mahdian. 2015. Incentives in large random two-sided markets. *ACM Transactions on Economics and Computation (TEAC)* 3, 3 (2015), 1–25.
[29] Ekhine Irurozki. 2014. *Sampling and learning distance-based probability models for permutation spaces*. Ph.D. Dissertation.
[30] Ekhine Irurozki, Borja Calvo, and Jose A Lozano. 2019. Mallows and generalized Mallows model for matchings. *Bernoulli* 25, 2 (2019), 1160–1188.
[31] Robert W Irving and Paul Leather. 1986. The complexity of counting stable marriages. *SIAM J. Comput.* 15, 3 (1986), 655–667.
[32] Robert W Irving, Paul Leather, and Dan Gusfield. 1987. An efficient algorithm for the "optimal" stable marriage. *Journal of the ACM (JACM)* 34, 3 (1987), 532–543.
[33] Robert W Irving, David F Manlove, and Sandy Scott. 2008. The stable marriage problem with master preference lists. *Discrete Applied Mathematics* 156, 15 (2008), 2959–2977.
[34] Will Jones and Alexander Teytelboym. 2017. Matching Systems for Refugees. *Journal on Migration and Human Security* 5, 3 (2017), 667–681. https://doi.org/10.1177/233150241700500306
[35] Yuichiro Kamada and Fuhito Kojima. 2015. Efficient matching under distributional constraints: Theory and applications. *American Economic Review* 105, 1 (2015), 67–99.
[36] Akiko Kato. 1993. Complexity of the sex-equal stable marriage problem. *Japan Journal of Industrial and Applied Mathematics* 10, 1 (1993), 1–19.
[37] Maurice G Kendall. 1938. A new measure of rank correlation. *Biometrika* 30, 1/2 (1938), 81–93.
[38] Bettina Klaus and Flip Klijn. 2006. Procedurally fair and stable matching. *Economic Theory* 27, 2 (2006), 431–447.
[39] Vicki Knoblauch. 2009. Marriage matching and gender satisfaction. *Social Choice and Welfare* 32, 1 (2009), 15–27.
[40] Donald E Knuth, Rajeev Motwani, and Boris Pittel. 1990. Stable husbands. *Random Structures & Algorithms* 1, 1 (1990), 1–14.
[41] Fuhito Kojima and Parag A Pathak. 2009. Incentives and stability in large two-sided matching markets. *American Economic Review* 99, 3 (2009), 608–27.
[42] SangMok Lee. 2016. Incentive compatibility of large centralized matching markets. *The Review of Economic Studies* 84, 1 (2016), 444–463.
[43] Avi William Levy. 2017. *Novel uses of the Mallows model in coloring and matching*. Ph.D. Dissertation. Dept. of Mathematics, University of Washington, Seattle.
[44] Tyler Lu and Craig Boutilier. 2014. Effective sampling and learning for mallows models with pairwise-preference data. *The Journal of Machine Learning Research* 15, 1 (2014), 3783–3829.
[45] Colin L Mallows. 1957. Non-null ranking models. I. *Biometrika* 44, 1/2 (1957), 114–130.
[46] Nicholas Mattei and Toby Walsh. 2013. PrefLib: A Library of Preference Data HTTP://PREFLIB.ORG. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT 2013) (Lecture Notes in Artificial Intelligence)*. Springer.
[47] Eric McDermid and Robert W Irving. 2014. Sex-equal stable matchings: Complexity and exact algorithms. *Algorithmica* 68, 3 (2014), 545–570.
[48] David G McVitie and Leslie B Wilson. 1971. The stable marriage problem. *Commun. ACM* 14, 7 (1971), 486–490.
[49] Morikazu Nakamura, Kenji Onaga, Seiki Kyan, and Manuel Silva. 1995. Genetic algorithm for sex-fair stable marriage problem. In *Proceedings of ISCAS'95-International Symposium on Circuits and Systems*, Vol. 1. IEEE, 509–512.
[50] Parag A Pathak and Tayfun Sönmez. 2008. Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review* 98, 4 (2008), 1636–52.
[51] Boris Pittel. 1989. The average number of stable matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 530–549.
[52] Ross Rheingans-Yoo and Jane Street. 2020. *Large random matching markets with localized preference structures can exhibit large cores*. Technical Report. Mimeo.
[53] Alvin E Roth. 1982. The economics of matching: Stability and incentives. *Mathematics of operations research* 7, 4 (1982), 617–628.
[54] Alvin E Roth and Elliott Peranson. 1999. The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American economic review* 89, 4 (1999), 748–780.
[55] Alvin E Roth and John H Vande Vate. 1990. Random paths to stability in two-sided matching. *Econometrica: Journal of the Econometric Society* (1990), 1475–1480.
[56] Hannu Salonen and Mikko AA Salonen. 2018. Mutually best matches. *Mathematical Social Sciences* 91 (2018), 42–50.
[57] Weiran Shen, Pingzhong Tang, and Yuan Deng. 2018. Coalition manipulation of gale-shapley algorithm. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.
[58] Mark Steyvers, Brent Miller, Pernille Hemmer, and Michael D Lee. 2009. The Wisdom of Crowds in the Recollection of Order Information. In *Advances in Neural Information Processing Systems*. 1785–1793.
[59] Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan. 2001. Gale-shapley stable marriage problem revisited: Strategic issues and applications. *Management Science* 47, 9 (2001), 1252–1267.
[60] Nikolaos Tziavelis, Ioannis Giannakopoulos, Katerina Doka, Nectarios Koziris, and Panagiotis Karras. 2019. Equitable Stable Matchings in Quadratic Time. In *Advances in Neural Information Processing Systems*. 455–465.
[61] Nikolaos Tziavelis, Ioannis Giannakopoulos, Rune Quist Johansen, Katerina Doka, Nectarios Koziris, and Panagiotis Karras. 2020. Fair Procedures for Fair Stable Marriage Outcomes. In *Thirty-Fourth AAAI Conference on Artificial Intelligence*.
[62] Rohit Vaish and Dinesh Garg. 2017. Manipulating Gale-Shapley Algorithm: Preserving Stability and Remaining Inconspicuous.. In *IJCAI*. 437–443.
[63] Hoang Huu Viet, SeungGwan Lee, and TaeChoong Chung. 2016. A bidirectional local search for the stable marriage problem. In *2016 International Conference on Advanced Computing and Applications (ACOMP)*. IEEE, 18–24.
[64] Hoang Huu Viet, Le Hong Trang, Le Pham Tuyen, and Taechoong Chung. 2020. A shortlist-based bidirectional local search for the stable marriage problem. *Journal of Experimental & Theoretical Artificial Intelligence* 32, 1 (2020), 147–163.