# A Map of Diverse Synthetic Stable Roommates Instances

Niclas Boehmer
Technische Universität Berlin
Berlin, Germany
niclas.boehmer@tu-berlin.de

Klaus Heeger
Technische Universität Berlin
Berlin, Germany
heeger@tu-berlin.de

Stanisław Szufa
AGH University
Kraków, Poland
szufa@agh.edu.pl

## ABSTRACT

Focusing on Stable Roommates (SR), we contribute to the toolbox for conducting experiments for stable matching problems. We introduce the polynomial-time computable mutual attraction distance to measure the similarity of SR instances, analyze its properties, and use it to create a map of SR instances. This map visualizes 460 synthetic SR instances (each sampled from one of ten different statistical cultures) as follows: Each instance is a point in the plane, and two points are close on the map if the corresponding SR instances are similar to each other. Subsequently, we conduct several exemplary experiments and depict their results on the map, illustrating the map's usefulness as a non-aggregate visualization tool, the diversity of our generated dataset, and the need to use instances sampled from different statistical cultures.

## KEYWORDS

stable matchings; framework for experiments; similarity of instances; statistical cultures

**ACM Reference Format:**
Niclas Boehmer, Klaus Heeger, and Stanisław Szufa. 2023. A Map of Diverse Synthetic Stable Roommates Instances. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 9 pages.

## 1 INTRODUCTION

Since their introduction by Gale and Shapley [22], stable matching problems have been extensively studied, both from a theoretical and a practical viewpoint. Numerous practical applications have been identified, and theoretical research has influenced the design of real-world matching systems [26, 31, 33]. In addition to the rich theoretical literature, there are also several works containing empirical investigations of stable matching problems (see [1, 8, 12–14, 16, 19, 23, 24, 27, 32, 34–38, 40, 43, 44] as a certainly incomplete list). Although these examples indicate that experimental works regularly occur, many papers on stable matchings do not include an experimental part and instead solely focus on the computational or axiomatic aspects of some mechanism or problem. However, to understand the properties of problems and mechanisms in practice, experiments are vital.

One reason for the lack of experimental work might be the rarity of real-world data (exceptions can be found in [16, 27, 34]). Consequently, researchers typically resort to some synthetic distribution, known as a statistical culture, for generating synthetic data. Remarkably, the vast majority of works simply use random preferences, where all possible valid preferences are sampled with the same probability (out of the nineteen works listed above, thirteen use this model, most of them as a single data source). However, as we will see later, instances with random preferences have very similar properties. Accordingly, conclusions drawn from experiments using only such instances (or, generally speaking, only instances sampled from one model) should be treated with caution, as it is unclear whether their results generalize.

With our work, we want to lay the foundation for more experimental work around stable matchings by introducing a measure for the similarity of instances and by creating a diverse synthetic dataset for testing together with a convenient framework to visualize and analyze it as a map (see Figure 1a for an example). We focus on instances of the Stable Roommates (SR) problem, where we have a set of agents, and each agent has strict preferences over all other agents. We selected the SR problem for this first, exemplary study because it is the mathematically most natural stable matching problem (agents' preferences do not contain ties and are complete, and there are no different "types" of agents). Consequently, statistical cultures for SR instances are relatively simple. Nevertheless, our general approach and several of our ideas and techniques can also be used to carry out similar studies for other stable matching problems, as demonstrated in Section 6, where we describe how to adapt our results to Stable Marriage (SM) instances (SM is the bipartite analogue of SR).

As part of our agenda to empower experimental work on stable matchings, we carry out the following steps:

*Distances Between SR Instances (Section 3).* To judge the diversity of a dataset for testing and to compare different statistical cultures to each other, a similarity measure is needed. We introduce the notion of isomorphism between SR instances and show how distances between preference orders naturally extend to distances between SR instances. Most importantly, we propose the polynomial-time computable mutual attraction distance[1], which we use in the following. To better understand the space of SR instances induced by our mutual attraction distance, we introduce four canonical "extreme" instances which are far away from each other.

*A Map of Synthetic SR Instances (Section 4).* We define a variety of statistical cultures to generate SR instances. From them, we generate a diverse test set for experimental work and picture it as a map of SR instances, a convenient framework to visualize non-aggregate experimental results. Moreover, we give intuitive interpretations of the different areas on the map. We also analyze how different statistical cultures relate to each other.

*Using the Map of SR Instances (Section 5).* To demonstrate possible use cases for the map, we perform exemplary experimental studies.

*Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.* 

[1] Note that we use the terms "distance (measure)" in an informal sense to refer to some function mapping pairs of instances to a positive real number; in particular, all our distance measures are pseudometrics but not all are metrics.

We analyze different quality measures for stable matchings, the number of blocking pairs for random matchings, and the running time to compute an “optimal" stable matching using an ILP. In sum, the instance-based view on experimental results provided by the map allows us to identify several interesting phenomena, for example, that instances sampled from the same culture all behave very similarly in our experiments. We further observe that instances from the same area of the map exhibit a similar behavior.

From a methodological perspective, our work follows a series of recent papers on (ordinal) elections [4, 6, 7, 20, 41]: Faliszewski et al. [20] introduced the problem of computing the distance between elections, focusing on isomorphic distances. Following up on this, Szufa et al. [41] created a dataset of synthetic elections sampled from a variety of different cultures and visualized them as a map of elections. Subsequently, Boehmer et al. [6] added several canonical elections to the map to give absolute positions a clearer meaning and added some real-world elections. Recently, Szufa et al. [42] created and analyzed a map of approval elections. The usefulness of the maps has already been demonstrated in different contexts. For example, Szufa et al. [41] identified that for elections from a certain region of the map, election winners are particularly hard to compute, Boehmer et al. [6] and Boehmer and Schaar [10] analyzed the nature and relationship of real-world elections by placing them on the map, and Boehmer et al. [5] evaluated the robustness of election winners using the map. Although our general agenda and approach are similar to the works of Faliszewski et al. [20], Szufa et al. [41] and Boehmer et al. [6], the intermediate steps, used distance measures, cultures, experiments, and technical details are naturally quite different.

The full version of this paper containing all proofs and additional discussions and experiments is available at arxiv.org/pdf/2208.04041.pdf [9]. The code for generating the map and conducting our experiments is available at https://github.com/szufix/mapel. The generated datasets of SR and SM instances are available at https://github.com/szufix/mapel_data.

## 2 PRELIMINARIES

Let $A$ be a finite set of agents. We denote by $\mathcal{L}(A)$ the set of all strict, total orders over $A$ which we call preference orders. We usually denote elements of $\mathcal{L}(A)$ as $\succ$ and for three agents $a$, $b$, and $c$, we say that $a$ is preferred to $b$ is preferred to $c$ if $a \succ b \succ c$. Moreover, for a preference order $\succ \in \mathcal{L}(A)$ and an agent $a \in A$, let $\mathrm{pos}_{\succ}(a)$ denote the position of $a$ in $\succ$, i.e., the number of agents that are preferred to $a$ in $\succ$ plus one. Furthermore, for $i \in [|A|]$, let $\mathrm{ag}_{\succ}(i)$ be the agent ranked in $i$-th position in $\succ$, i.e., the agent $b \in A$ such that $i = \mathrm{pos}_{\succ}(b)$.

For two preference orders $\succ, \succ' \in \mathcal{L}(A)$, their swap distance $\mathrm{swap}(\succ, \succ')$ is the number of agent pairs on whose ordering $\succ$ and $\succ'$ disagree. For two preference orders $\succ, \succ' \in \mathcal{L}(A)$, their Spearman distance $\mathrm{spear}(\succ,\succ')$ is $\sum_{a \in A} |\mathrm{pos}_{\succ}(a) - \mathrm{pos}_{\succ'}(a)|$. As proven by Diaconis and Graham [17], it holds that $\mathrm{swap}(\succ, \succ') \leq \mathrm{spear}(\succ,\succ') \leq 2 \cdot \mathrm{swap}(\succ, \succ')$.

A Stable Roommates (SR) instance $\mathcal{I}$ consists of a set $A$ of agents, with each agent $a \in A$ having a preference order $\succ_a \in \mathcal{L}(A \setminus \{a\})$ over all other agents. For simplicity, we will focus on instances with an even number of agents.

A matching of agents $A$ is a subset of agent pairs $\{a, a'\}$ with $a \neq a' \in A$ where each agent appears in at most one pair. We say that an agent is unmatched in a matching $M$ if $a$ does not appear in any pair from $M$; otherwise, we say that $a$ is matched. For a matched agent $a \in A$ and a matching $M$, we write $M(a)$ to denote the partner of $a$ in $M$, i.e., $M(a) = a'$ if $\{a, a'\} \in M$. A pair $\{a, a'\}$ of agents blocks a matching $M$ if $a$ is unmatched or prefers $a'$ to $M(a)$ and $a'$ is unmatched or prefers $a$ to $M(a')$. A matching that is not blocked by any agent pair is called a stable matching.

For two sets $X$ and $Y$ with $|X| = |Y|$, we denote by $\Pi(X, Y)$ the set of all bijections $\sigma : X \rightarrow Y$ between $X$ and $Y$. Let $A$ and $A'$ be two sets of agents with $|A| = |A'|$ and let $\sigma \in \Pi(A, A')$. Then, for an agent $a \in A$ and a preference order $\succ_a \in \mathcal{L}(A \setminus \{a\})$, we write $\sigma(\succ_a)$ to denote the preference order over $A' \setminus \{\sigma(a)\}$ arising from $\succ_a$ by replacing each agent $b \in A \setminus \{a\}$ by $\sigma(b) \in A' \setminus \{\sigma(a)\}$.

## 3 DISTANCE MEASURES

This section is devoted to measuring the distance between two SR instances, a key ingredient of our map. Other use cases include meaningful selecting test instances, comparing different statistical cultures, and analyzing real-world instances. Specifically, in Section 3.1, we define an isomorphism between two SR instances, show how distance measures over preference orders can be generalized to distance measures over SR instances, and prove that computing the Spearman distance between SR instances is computationally intractable. In Section 3.2, we introduce our mutual attraction distance and make some observations concerning its properties and the associated mutual attraction matrices.

### 3.1 Isomorphism and Isomorphic Distances

Two SR instances are isomorphic if renaming the agents in one instance can produce the other. For this, as each agent is associated with a preference order defined over other agents, a single mapping suffices. Accordingly, we define an isomorphism on SR instances:

Definition 1. *Two SR instances $(A, (\succ_a)_{a \in A})$ and $(A', (\succ_{a'})_{a' \in A'})$ with $|A| = |A'|$ are isomorphic if there is a bijection $\sigma : A \rightarrow A'$ such that $\succ_{\sigma(a)} = \sigma(\succ_a)$ for all $a \in A$.*

We give an example for two isomorphic SR instances:

Example 2. *Let $\mathcal{I}$ with agents $a, b, c$, and $d$ and $\mathcal{I}'$ with agents $x, y, z$, and $w$ be two SR instances with the following preferences:*

$$a : b \succ c \succ d, \quad b : c \succ a \succ d, \quad c : b \succ d \succ a, \quad d : a \succ c \succ b,$$
$$x : y \succ w \succ z, \quad y : z \succ w \succ x, \quad z : w \succ y \succ x, \quad w : z \succ x \succ y.$$

*$\mathcal{I}$ and $\mathcal{I}'$ are isomorphic as witnessed by the mapping $\sigma(a) = y$, $\sigma(b) = z$, $\sigma(c) = w$, and $\sigma(d) = x$.*

One can easily check whether two SR instances $(A, (\succ_a)_{a \in A})$ and $(A', (\succ_{a'})_{a' \in A'})$ are isomorphic: Assuming that an isomorphism $\sigma_{a'}$ maps $a \in A$ to $a' \in A'$, then this already completely characterizes $\sigma_{a'}$, as for any $b \in A \setminus \{a\}$ with $\mathrm{pos}_{\succ_a}(b) = i$, we must have $\sigma_{a'}(b) = \mathrm{ag}_{\succ'_{a'}}(i)$. Thus, it suffices to fix an arbitrary agent $a \in A$ and then check for each $a' \in A'$ whether $\sigma_{a'}$ is an isomorphism.

Observation 3. *Deciding whether two SR instances with $2n$ agents are isomorphic can be done in $O(n^3)$ time.*

For any distance measure $p$ between preference orders, our notion of isomorphism can be easily used to extend $p$ to a distance measure over SR instances: The resulting distance between two SR instances is the minimum (over all bijections $\sigma$ between the agent sets) sum (over all agents) of the distance between the preferences of $a$ and the preferences of $\sigma(a)$ (measured by $p$):

Definition 4. *Let $p$ be a distance measure between preference orders. Let $\mathcal{I} = (A, (\succ_a)_{a\in A})$ and $\mathcal{I}' = (A', (\succ_{a'})_{a'\in A'})$ be two SR instances with $|A| = |A'|$. Their $d_p$ distance is: $d_p(\mathcal{I}, \mathcal{I}') := \min_{\sigma\in\Pi(A,A')} \sum_{a\in A} p(\sigma(\succ_a), \succ_{\sigma(a)})$.*

For all distance measures $p$ between preference orders where $p(x, y) = 0$ if and only if $x = y$, for any two SR instances $\mathcal{I}$ and $\mathcal{I}'$ it holds that $d_p(\mathcal{I}, \mathcal{I}') = 0$ if and only if $\mathcal{I}$ and $\mathcal{I}'$ are isomorphic. We will also call such a distance an *isomorphic distance.*

Example 5. *Applying Definition 4, the Spearman distance* spear$(\cdot,\cdot)$ *and the swap distance* swap$(\cdot,\cdot)$ *between preference orders (as defined in Section 2) can be lifted to distance measures $d_{\text{spear}}$ and $d_{\text{swap}}$ between SR instances. Let $\mathcal{I}$ with agents $a, b, c$, and $d$ and $\mathcal{I}'$ with agents $x, y, z$, and $w$ be two SR instances with the following preferences:*

$$a : b \succ c \succ d, \quad b : a \succ c \succ d, \quad c : a \succ b \succ d, \quad d : a \succ b \succ c,$$
$$x : y \succ z \succ w, \quad y : x \succ z \succ w, \quad z : w \succ y \succ x, \quad w : z \succ y \succ x.$$

*Then, for the mapping $\sigma(a) = x$, $\sigma(b) = y$, $\sigma(c) = z$, and $\sigma(d) = w$, the Spearman distance of $\mathcal{I}$ and $\mathcal{I}'$ is 8 and the swap distance is 6. While for the Spearman distance this is the optimal mapping (so $d_{\text{spear}}(\mathcal{I}, \mathcal{I}') = 8$) for the swap distance the mapping $\sigma(a) = y$, $\sigma(b) = x$, $\sigma(c) = z$, and $\sigma(d) = w$ results in a smaller distance of 4. Indeed, we have $d_{\text{swap}}(\mathcal{I}, \mathcal{I}') = 4$.*

We consider the Spearman distance $d_{\text{spear}}$ and the swap distance $d_{\text{swap}}$ as "ideal" distances, as they are quite fine-grained and *isomorphic* distances. Unfortunately, both are hard to compute. For $d_{\text{swap}}$, this follows from the NP-hardness of computing the Kemeny score of an election [18]. Moreover, we show that computing the Spearman distance between two SR instances is at least as hard as deciding whether two graphs are isomorphic, which is a famous candidate for the complexity class NP-intermediate.

Proposition 6. *There is no polynomial-time algorithm to compute $d_{\text{spear}}$, unless Graph Isomorphism is in P.*

## 3.2 Mutual Attraction Distance

In this section, we introduce and discuss our main distance measure, which we call mutual attraction distance.

*3.2.1 Intuition.* One characteristic of SR instances is that each agent is associated with a preference order and also appears in the preference order of other agents. Thus, when considering, for instance, stable matchings, for an agent $a$ it is not only important which agents $a$ likes, but also whether they like $a$ as well. Accordingly, our mutual attraction distance focuses on how pairs of agents rank each other. In particular, each agent $a$ is characterized by a mutual attraction vector whose $i$-th entry contains the position in which $a$ appears in the preferences of the agent who $a$ ranks in $i$-th position. In the mutual attraction distance, we match the agents from two different instances such that the $\ell_1$ distance between the mutual attraction vectors of matched agents is minimized.

*3.2.2 Notation.* For $p, q \in \mathbb{N}$, some $i \in [p]$, and a matrix $M \in \mathbb{N}^{p\times q}$, let $M_i$ denote the $i$-th row of $M$. For an SR instance $\mathcal{I} = (A = \{a_1, \dots a_{2n}\}, (\succ_a)_{a\in A})$, an agent $a \in A$, and some $i \in [2n-1]$, let $\mathcal{MA}^{\mathcal{I}}(a, i)$ be the position of $a$ in the preference order of the agent $a'$ which is ranked in position $i$ by $a$, i.e., $\mathcal{MA}^{\mathcal{I}}(a, i) := \text{pos}_{\succ_{a'}}(a)$ where $a' := \text{ag}_{\succ_a}(i)$. The mutual attraction vector of agent $a$ is $\mathcal{MA}^{\mathcal{I}}(a) = \big(\mathcal{MA}^{\mathcal{I}}(a, 1), \dots, \mathcal{MA}^{\mathcal{I}}(a, 2n-1)\big)$. Lastly, the mutual attraction matrix $\mathcal{MA}^{\mathcal{I}}$ of $\mathcal{I}$ is the matrix whose $i$-th row is the vector $\mathcal{MA}(a_i)^{\mathcal{I}}$.

Definition 7. *For two mutual attraction matrices of SR instances $\mathcal{I}$ and $\mathcal{I}'$ on $2n$ agents, we define their* mutual attraction distance *as*

$$\mathrm{d}_{\mathrm{MAD}}(\mathcal{MA}^{\mathcal{I}}, \mathcal{MA}^{\mathcal{I}'}) := \min_{\sigma\in\Pi([2n],[2n])} \sum_{i\in[2n]} \ell_1(\mathcal{MA}^{\mathcal{I}}_i, \mathcal{MA}^{\mathcal{I}'}_{\sigma(i)}).$$

*The* mutual attraction distance $\mathrm{d}_{\mathrm{MAD}}(\mathcal{I}, \mathcal{I}')$ *between two SR instances $\mathcal{I}$ with agents $A$ and $\mathcal{I}'$ with agents $A'$ with $|A| = |A'|$ is the mutual attraction distance of their mutual attraction matrices.*

Example 8. *Consider the two SR instances $\mathcal{I}$ and $\mathcal{I}'$ defined in Example 5. Their mutual attraction matrices are:*

$$\mathcal{MA}^{\mathcal{I}} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix} \end{array}, \quad \mathcal{MA}^{\mathcal{I}'} = \begin{array}{c} \\ x \\ y \\ z \\ w \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 3 & 3 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} \end{array}$$

*Their mutual attraction distance is $2 + 0 + 2 + 2 = 6$ as witnessed by the mapping $\sigma(a) = z$, $\sigma(b) = y$, $\sigma(c) = x$, and $\sigma(d) = w$.*

*3.2.3 Computation.* Given two SR instances $\mathcal{I}$ over agents $A$ and $\mathcal{I}'$ over agents $A'$ with $|A| = |A'|$, computing their mutual attraction distance reduces to finding a minimum-weight perfect matching in a complete bipartite graph $G = (A \cup A', E)$ where edge $\{a, a'\} \in E$ has weight $\ell_1(\mathcal{MA}^{\mathcal{I}}(a), \mathcal{MA}^{\mathcal{I}'}(a'))$.

Observation 9. *Given two SR instances $\mathcal{I}$ and $\mathcal{I}'$ with $2n$ agents each,* $\mathrm{d}_{\mathrm{MAD}}(\mathcal{I}, \mathcal{I}')$ *can be computed in $O(n^3)$ time.*

*3.2.4 Realizable Mutual Attraction Matrices.* Not every $(2n) \times (2n-1)$-matrix is the mutual attraction matrix of some SR instance. Accordingly, we call a matrix $M$ *realizable* if there is an SR instance $\mathcal{I}$ with $\mathcal{MA}^{\mathcal{I}} = M$. Realizable matrices exhibit certain characteristics. For example, since each agent ranks exactly one agent at position $j$ for every $j \in [2n-1]$, every realizable matrix $M \in \mathbb{N}^{(2n)\times(2n-1)}$ contains each number from $[2n-1]$ exactly $2n$ times. Unfortunately, checking whether a matrix is realizable is NP-hard.

Theorem 10. *Given a $(2n) \times (2n-1)$ matrix $M$, deciding if there is an SR instance $\mathcal{I}$ with $\mathcal{MA}^{\mathcal{I}} = M$ is NP-complete.*

*3.2.5 Isomorphism Property.* Unfortunately, in contrast to the swap and Spearman distance, the mutual attraction distance is not isomorphic (see our full version [9] for formal statements and proofs).[2]

[2]Note that the positionwise distance used in the ”map of elections“ framework is also not isomorphic [4, 6, 7, 20, 41].

In fact, there even exist mutual attraction matrices realized by two non-isomorphic SR instances $\mathcal{I}_1$ and $\mathcal{I}_2$ where $\mathcal{I}_1$ admits a stable matching but $\mathcal{I}_2$ does not. This indicates that the mutual attraction distance between two instances has only a limited predictive value for their relationship in terms of their (distance to) stability, which not too surprising given that stability is dependent on local configurations. However, our synthetically generated instances are anyway close to admitting a stable matching in the sense that in all instances there is a matching blocked by only a few pairs.

*3.2.6 Correlation With Spearman Distance.* Next, we consider the correlation between the mutual attraction distance and the appealing yet hard to compute Spearman distance. While we prove that the ratio between the two can be unbounded, in practice the two are highly correlated: On the test dataset of 460 instances described in Section 4.1 for twelve agents (which is the largest number of agents for which we could compute the Spearman distance via a brute-force approach in weeks), the Pearson Correlation Coefficient (PCC) between the mutual attraction and Spearman distances is 0.801, which is typically regarded as a strong correlation [39]. In particular, for 95% of instance pairs $(\mathcal{I}, \mathcal{I}')$ we have that $0.82 \cdot \mathrm{d_{MAD}}(\mathcal{I}, \mathcal{I}') \leq d_{\mathrm{spear}}(\mathcal{I}, \mathcal{I}') \leq 1.48 \cdot \mathrm{d_{MAD}}(\mathcal{I}, \mathcal{I}')$.

*3.2.7 Navigating the Space of SR Instances.* Interpreting a map of SR instances, it will be useful to give different regions on the map an intuitive meaning. This is why we now identify four somewhat "canonical" extreme mutual attraction matrices falling into four very different parts of the map.

*Identity.* Our first extreme case is that all agents have the same preferences (also called master list), a setting which has already attracted significant attention in the literature [11, 15, 28, 30]. For $n \in \mathbb{N}$, for each $i \in [2n]$ and $j \in [2n-1]$, the identity matrix is $\mathrm{ID}^{2n}[i,j] = i$ if $j \geq i$ and $\mathrm{ID}^{2n}[i,j] = i-1$ if $j < i$ .

*Mutual Agreement.* Our second extreme case is mutual agreement: For each pair $a$ and $a'$ of agents, $a$ and $a'$ evaluate each other identically, i.e., $a$ ranks $a'$ on the $i$-th position if and only if $a'$ ranks $a$ on the $i$-th position. For $n \in \mathbb{N}$, this is captured in the mutual agreement matrix $\mathrm{MA}^{2n}$ where we have $\mathrm{MA}^{2n}[i,j] = j$ for each $i \in [2n]$ and $j \in [2n-1]$.

*Mutual Disagreement.* Our third extreme case is mutual disagreement. For each pair $a$ and $a'$ of agents, their evaluations for each other are diametrical, i.e., $a$ ranks $a'$ in the $i$-th position if and only if $a'$ ranks $a$ in the $(2n-i)$-th position. For $n \in \mathbb{N}$, this is captured in the mutual disagreement matrix $\mathrm{MD}^{2n}$ where we have $\mathrm{MD}^{2n}[i,j] = 2n-j$ for each $i \in [2n]$ and $j \in [2n-1]$.

*Chaos.* Our fourth extreme mutual attraction matrix is the chaos matrix $\mathrm{CH}^{2n}$, which is defined for each $i \in [2n]$ and $j \in [2n-1]$ as $\mathrm{CH}^{2n}[i,j] = j$ if $i = 1$ and $\mathrm{CH}^{2n}[i,j] = i + nj - n - 1 \mod 2n-1$ if $i > 1$. We have no natural interpretation of the chaos matrix. We added this matrix because it is far away from the other three and thus falls into an otherwise vacant part of the map. Its name "chaos" stems from the fact that this matrix is close on the map to instances with uniformly at random sampled preferences.

In our full version [9], we prove that the first three matrices are always realizable, while $\mathrm{CH}^{2n}$ is realizable for all $n$ where $2n-1$ is not divisible by 3. Moreover, we show that the maximum mutual attraction distance between two SR instances with $2n$ agents is $D(2n) := 4 \cdot (n-1) \cdot n^2$. We further prove that the mutual agreement matrix and the mutual disagreement matrix match this bound, therefore forming a diameter of our space. For each two matrices $X$ and $Y$ among ID, MA, CH, and MD, we define their asymptotic normalized distance as $\mathrm{nd_{MAD}}(X,Y) := \lim_{n\to\infty} \mathrm{d_{MAD}}(X^{2n}, Y^{2n})/D(2n)$. We further prove that for all pairs of matrices $X, Y \in \{\mathrm{ID}, \mathrm{MA}, \mathrm{MD}, \mathrm{CH}\}$ with $\{X, Y\} \neq \{\mathrm{MA}, \mathrm{MD}\}$ we have $\mathrm{nd_{MAD}}(X,Y) = \frac{2}{3}$. This implies that our extreme matrices are indeed far from each other. In the following, all mentioned values of the mutual attraction distance are normalized values, i.e., they are divided by $D(2n)$.

# 4 A MAP OF SYNTHETIC SR INSTANCES

We present a map of synthetic SR instances. In Section 4.1, we describe how we create the map and how we generate the instances. In Section 4.2, we explain the map by giving the horizontal and vertical axis a natural interpretation and analyzing where different statistical cultures land.

## 4.1 Creating the Map

We first describe our dataset of 460 SR instances.

*4.1.1 Points on the Map – Statistical Cultures.* We use the following statistical cultures. To the best of our knowledge, only the Impartial Culture, Attributes, Mallows, and Euclidean models have been previously considered.

**Impartial Culture (IC).** Agent $a \in A$ draws its preferences uniformly at random from $\mathcal{L}(A \setminus \{a\})$.

**2-IC.** Given some $p \in [0, 0.5]$, we partition $A$ into two sets $A_1 \cupdot A_2$ with $|A_1| = \lfloor p \cdot |A| \rfloor$. Each agent $a \in A$ samples a preference order $\succ$ from $\mathcal{L}(A_1 \setminus \{a\})$ and $\succ'$ from $\mathcal{L}(A_2 \setminus \{a\})$. If $a \in A_1$, then $a$'s preferences start with all agents from $A_1$ ordered according to $\succ$ and then all agents from $A_2$ ordered according to $\succ'$. If $a \in A_2$, then it is the other way around, i.e., the preferences start with $\succ'$ and end with $\succ$. The intuition is that there are two groups of different sizes (e.g., representing demographic groups), and each agent prefers all agents from its group to agents from the other group, but preferences within the group are random.

**Mallows.** For a dispersion parameter $\phi \in [0,1]$ and a preference order $\succ^* \in \mathcal{L}(A)$, the Mallows distribution $\mathcal{D}^{\succ^*,\phi}_{\mathrm{Mallows}}$ assigns preference order $\succ \in \mathcal{L}(A)$ a probability proportional to $\phi^{\mathrm{swap}(\succ^*,\succ)}$ (for $\phi = 1$, we get IC).[3] To sample an SR instance, given a dispersion parameter $\phi \in [0,1]$, we draw $\succ^*$ uniformly at random from $\mathcal{L}(A)$. For each agent $a \in A$, we then obtain its preferences by drawing a preference order from $\mathcal{D}^{\succ^*,\phi}_{\mathrm{Mallows}}$ and deleting $a$. The intuition is that there is a ground truth and agents have a given likelihood to deviate from it.

**Euclidean.** [2]. Given some $d \in \mathbb{N}$, for each agent $a \in A$, we uniformly at random sample a point $\mathbf{p}^a$ from $[0,1]^d$. Agent $a$ ranks other agents increasingly by the Euclidean distance between their points, i.e., by $\ell_2(\mathbf{p}^a, \mathbf{p}^b)$ for $b \in A \setminus \{a\}$. The intuition is that each

[3] In fact, we use a normalized variant of Mallows model [6] with a normalized dispersion parameter norm-$\phi$, which is internally converted to a value of $\phi$. The higher norm-$\phi$, the higher is the expected swap distance of a sampled preference order to $\succ^*$. Due to the normalization, for orders sampled for norm-$\phi = 0.5$, the expected swap distance from $\succ^*$ is half of the expected distance for orders sampled for norm-$\phi = 1$.

dimension represents some continuous property of the agents, and agents prefer similar agents.

**Reverse-Euclidean.** Given some $p \in [0,1]$ and $d \in \mathbb{N}$, we partition $A$ into two sets $A_1 \cup A_2$ with $|A_1| = \lfloor p \cdot |A| \rfloor$. Again, each agent corresponds to some uniformly at random sampled point $\mathbf{p}^a$ from $[0,1]^d$ and ranks other agents according to their Euclidean distance. However, here an agent $a \in A_1$ ranks agents decreasingly by their Euclidean distance to $\mathbf{p}^a$ and an agent $a \in A_2$ ranks agents increasingly by their Euclidean distance to $\mathbf{p}^a$. The intuition is similar to Euclidean, but a $p$-fraction of agents prefer agents that are different from them.

**Mallows-Euclidean.** Given a normalized dispersion parameter norm-$\phi \in [0,1]$ and some $d \in \mathbb{N}$, we start by generating agents' intermediate preferences $(\succ_a)_{a \in A}$ according to the Euclidean model with $d$ dimensions. Subsequently, for each $a \in A$, we obtain its final preferences by sampling a preference order from $\mathcal{D}_{\text{Mallows}}^{\succ_a, \text{norm-}\phi}$. The resulting instances are perturbed Euclidean instances.

**Expectations-Euclidean.** Given some $d \in \mathbb{N}$ and $\sigma \in \mathbb{R}^+$, for each agent $a \in A$, we sample one point $\mathbf{p}^a$ uniformly at random from $[0,1]^d$. Subsequently, we sample a second point $\mathbf{q}^a$ from $[0,1]^d$ using a $d$-dimensional Gaussian function with mean $\mathbf{p}^a$ and standard deviation $\sigma$. Agent $a$ ranks the agents increasingly according to $\ell_2(\mathbf{p}^a, \mathbf{q}^b)$ for $b \in A \setminus \{a\}$. Again, agents are characterized by continuous attributes; however, their "ideal" points are not necessarily where they are.

**Fame-Euclidean.** Given some $d \in \mathbb{N}$ and $f \in [0,1]$, we sample for each agent $a \in A$ uniformly at random a point $\mathbf{p}^a \in [0,1]^d$ and a number $f^a \in [0, f]$. Agent $a$ ranks the other agents increasingly by $\ell_2(\mathbf{p}^a, \mathbf{p}^b) - f^b$ for $b \in A \setminus \{a\}$. The intuition is similar as for Euclidean, but some agents have a higher quality/fame $f^a$ and are thus more attractive to everyone.

**Attributes.** [3]. Given some $d \in \mathbb{N}$, for each agent $a \in A$ we uniformly at random sample $\mathbf{p}^a \in [0,1]^d$ and $\mathbf{w}^a \in [0,1]^d$. Agent $a$ ranks the other agents decreasingly by the inner product of $\mathbf{w}^a$ and $\mathbf{p}^b$, i.e., by $\sum_{i \in [d]} \mathbf{w}_i^a \cdot \mathbf{p}_i^b$. The intuition is that there are different objective evaluation criteria and agents assign different importance to them.

**Mallows-MD.** Given a normalized dispersion parameter norm-$\phi \in [0,1]$, we start with an instance that realizes the mutual disagreement matrix $\mathrm{MD}^{2n}$ where for each $i \in [2n]$ agent $a_i$ has preferences $a_{i+1} \succ_{a_i} a_{i+2} \succ_{a_i} \cdots \succ_{a_i} a_n \succ_{a_i} a_1 \succ_{a_i} a_2 \succ_{a_i} \cdots \succ_{a_i} a_{i-1}$. Subsequently, for each $a_i \in A$, we obtain its final preferences by sampling a preference order from $\mathcal{D}_{\text{Mallows}}^{\succ_{a_i}, \text{norm-}\phi}$. The reason we consider this model is that it covers a part of the map that would otherwise remain uncovered.

Our dataset consists of 460 instances sampled from the above-described statistical cultures. That is, we sampled 20 instances for each of the following cultures: Impartial Culture, 2-IC with $p \in \{0.25, 0.5\}$, Mallows with norm-$\phi \in \{0.2, 0.4, 0.6, 0.8\}$, 1D and 2D Euclidean, Reverse-Euclidean with $d = 2$ and $p \in \{0.05, 0.15, 0.25\}$, Mallows-Euclidean with $d = 2$ and norm-$\phi \in \{0.2, 0.4\}$, Expectations-Euclidean with $d = 2$ and $\sigma \in \{0.2, 0.4\}$, Fame-Euclidean with $d = 2$ and $f \in \{0.2, 0.4\}$, Attributes with $d \in \{2, 5\}$, and Mallows-MD with norm-$\phi \in \{0.2, 0.4, 0.6\}$. In addition, on our maps, we include the four extreme matrices described in Section 3.2.7.

*4.1.2 Drawing the Map.* To draw a map of our dataset, we compute for each pair of instances their mutual attraction distance. Subsequently, we embed the instances as points in the two-dimensional Euclidean space. Our goal is that the Euclidean distance of two points reflects the mutual attraction distance between the two respective instances. To obtain the embedding, we use a variant of the force-directed algorithm of Kamada and Kawai [29].[4] We depict the map visualizing our dataset of 460 instances for 200 agents in Figure 1a. [5]

To correctly interpret the map, we stress that our embedding algorithm does not optimize some global objective function. Instead, the algorithm works in a decentralized fashion also aiming at producing a visually pleasing image. Consequently, the position of instances on the map can be different in different runs and certainly depend on which other instances are part of the map. To verify the quality of the embedding, we computed the embedding's distortion and find that while the embedding is certainly not perfect, most of the distances are represented adequately. We want to remark that some error is to be expected here as the space of SR instances under the mutual attraction distance is highly complex; however, the general picture the map provides is indeed correct and helpful to get an intuitive interpretation of experimental results.

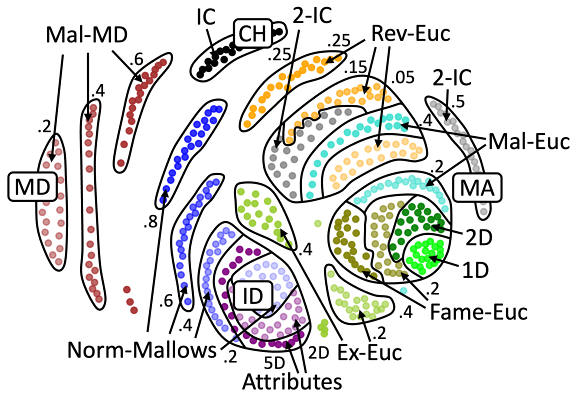
## 4.2 Understanding the Map

We now take a closer look at the map of SR instances shown in Figure 1a. Examining the map, what stands out is that for all cultures, instances sampled from this culture are placed close to each other on the map, resulting in an island-like structure. In fact, instances sampled from the same culture are usually close to each other under the mutual attraction distance (or at least closer to each other than to instances sampled from other cultures). While this is to be expected to a certain extent, this observation validates our approach in that the mutual attraction distance is seemingly able to identify the shared structure of instances sampled from the same statistical culture and in that our embedding algorithm is able to detect these clusters.

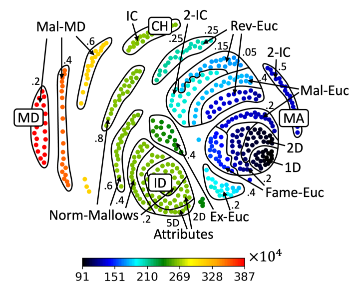
Moreover, interestingly, the different statistical cultures have a different "variation", i.e., the average mutual attraction distance of two instances sampled from the same culture substantially differs for the different cultures. The Impartial Culture model has with 0.59 the highest variation, while the Euclidean model for $d = 1$ has with 0.07 the lowest variation. The value for Impartial Culture is quite remarkable, as it means that Impartial Culture instances are on average almost as far away from each other as, for example, ID from the other extreme points. Because of the limitations of two-dimensional Euclidean space, this is not adequately represented on the map, as Impartial Culture instances are still placed close to each other. The reason for this is that they are all at a similar (even larger) distance to the other instances. In the following experiments,

[4] The algorithm starts with an arbitrary embedding of the instances. Then, it adds an attractive force between each pair of instances whose strength reflects their mutual attraction distance and a repulsive force between each pair ensuring that there is a certain minimum distance between each two points. Subsequently, the instances move based on the applied forces until a minimal energy state is reached. Szufa et al. [41] and Boehmer et al. [6] used the closely related Fruchterman–Reingold algorithm.

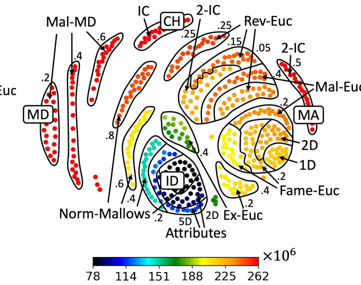
[5] We focus on 200 agents. Maps for different numbers of agents are available in the full version [9] and look quite similar.



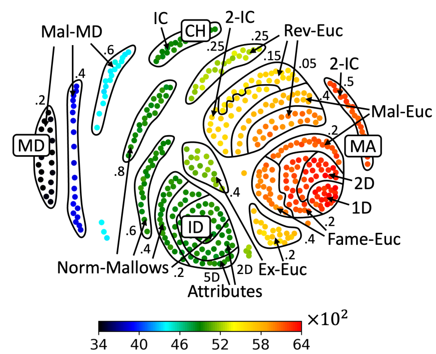

**(a) Colors indicate the statistical culture points were sampled from**



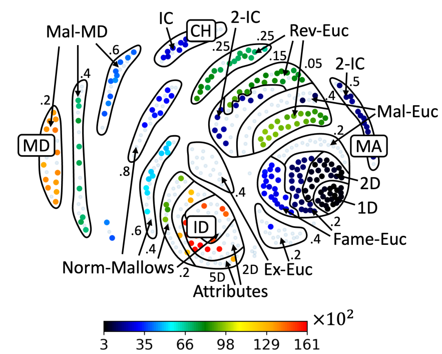

**(b) Mutuality**



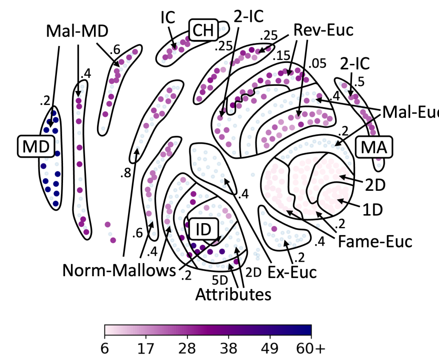

**(c) Rank distortion**




**(d) Average number of blocking pairs for a perfect matching**




**(e) Minimum summed rank of all agents for their partner in any stable matching**




**(f) Seconds needed to compute summed rank minimal stable matching**

**Figure 1: Map of 460 SR instances for 200 agents visualizing different quantities for each instance. Each instance is represented by a point. Roughly speaking, the closer two points are on the map, the more similar are the respective SR instances under the mutual attraction distance. Transparent points have no stable matching.**

we observe that Impartial Culture instances nevertheless behave quite similarly.

Taking a closer look at the map, we observe that our four extreme points fall into four different parts. On the right, we have the mutual agreement matrix MA. Accordingly, models for which mutual agreement is likely to appear all land in the right part of the map, namely, Euclidean instances (where intuitively speaking agent $a$ likes agent $b$ if they are close to each other making it also likely that $b$ likes $a$), the Fame-Euclidean model for $f = 0.2$, the Mallows-Euclidean model for norm-$\phi = 0.2$, the Reverse-Euclidean model for $p = 0.05$ (these three are all basically differently perturbed variants of Euclidean models and consequently also on average slightly further away from MA than Euclidean instances), and the 2-IC model for $p = 0.5$ (where we have some guaranteed level of mutual agreement because there are two groups of agents and agents from one group prefer each other to the agents from the other group).

On the left, we have the mutual disagreement matrix MD with only instances from the Mallows-MD model being close to it (in general, it is to be expected that if we apply the Mallows model on top of some other model $\mathcal{X}$, then for small values of norm-$\phi$ the sampled instances are close to the ones from $\mathcal{X}$ but move further and further away towards Impartial Culture instances as norm-$\phi$ grows). That the mutual (dis)agreement matrices are at the two ends of the horizontal axis raises the question whether the horizontal axis can be indeed interpreted as an indicator for the degree of mutuality in SR instances. This hypothesis gets strongly confirmed in Figure 1b where we color the points on the map according to their mutuality value, which we define as the total difference between the mutual evaluations of agent pairs, i.e., $\sum_{a \in A} \sum_{i \in [|A|-1]} |\mathcal{MA}(a,i) - i|$. The nicely continuous shading in Figure 1b indicates a strong correlation between the mutuality value of an instance and its $x$-coordinate on the map. Moreover, instances that are close on the map have indeed similar mutuality values. Moreover, the continuous coloring indicates that our dataset provides a good and almost uniform coverage of the space of SR instances (at least in terms of their mutuality value).

Turning to the middle part of the map, the identity matrix ID can be found at the bottom. Close to identity are instances from cultures where agent's quality is "objective". Namely, Mallows model with norm-$\phi = 0.2$ (where the preferences of agents are still often close to the central order) and the Attributes model with $d = 2$ (where each agent has two quality scores and the preferences of agents only differ in how they weight the quality scores). The chaos matrix CH is placed in the top part of the map together with the "chaotic" Impartial Culture instances. Mallows instances naturally form a continuous spectrum between identity and chaos. These observations give rise to the hypothesis that in instances placed at the bottom of the map most agents have similar preferences, while in instances placed at the top all agents have roughly the same quality and few agents are particularly (un)popular. To quantify whether agents agree or disagree on the quality of the agents, we measure the rank distortion of an instance, i.e., for each agent we sum up the absolute difference between all pairs of entries in its mutual attraction vector $\sum_{a \in A} \sum_{i,j \in [|A|-1]} |\mathcal{MA}(a,i) - \mathcal{MA}(a,j)|$. Note that, for example, for an agent that is always ranked in the same position by all other agents this absolute difference is zero. We show in Figure 1c a map colored by the rank distortion of instances. The picture here is slightly different than for the horizontal axis in that instances with the same $y$ coordinate might still have a very different rank distortion. In fact, what we rather see here is that the further a point is from ID on the map, the larger is its rank distortion and thus the higher is the disagreement concerning agents quality (which is quite intuitive recalling that for both MA and MD the rank distortion is maximal).

## 5 USING THE MAP

To illustrate the usefulness of the map to evaluate experiments and to confirm our previous observation that instances that are close to each other on the map have similar properties, we perform some exemplary experiments.

*Expected Number of Blocking Pairs.* Motivated by the fundamental importance of blocking pairs for stable matchings, we measure the expected number of blocking pairs for an arbitrary perfect matching. For this, for each instance, we sampled 100 perfect matchings uniformly at random and for each counted the number of blocking pairs. The results are depicted in Figure 1d. As for the mutuality value, we get a nicely continuous shading along the horizontal axis. This clear correlation with the mutuality value is quite intuitive as in case there is a high mutual agreement agents are also more likely to form blocking pairs (if an agent $a$ prefers an agent $b$ to its current partner, then because the mutuality is high $b$ also tends to like $a$ and tends to prefer $a$ to its current partner); if there is mutual disagreement, the picture is reversed (an agent prefers the agents to its current partner that tend to dislike it). This is also clearly visible in Figure 1d, as instances close to MD have a low expected number of blocking pairs, whereas for instances close to MA the expected number is much higher.

*Existence of a Stable Matching.* A fundamental question is which of our 460 instances admit a stable matching. In Figures 1e and 1f, if an instance admits a stable matching, it is drawn as a solid point and otherwise as a transparent point. Examining the map, we do not see a clear correlation between whether instances admit a stable matching and their position on the map. This is also quite intuitive, given that the existence of a stable matching might depend on some local configuration. However, what is clearly visible is that for different cultures the probability of admitting a stable matching is quite different: On the one hand, instances sampled from the Euclidean, Fame-Euclidean, and Reverse-Euclidean models almost always admit a stable matching (for the Euclidean model this is even guaranteed). On the other hand, instances sampled from the Mallows-Euclidean and Expectations-Euclidean model only very rarely admit a stable matching. The drastic contrast between the Euclidean model and the Mallows-Euclidean model with norm-$\phi = 0.2$ and between the Reverse-Euclidean and Expectations-Euclidean model here is quite remarkable, as they are conceptually similar. We also computed for each instance the minimum possible number of blocking pairs in a matching. Interestingly, this value has a stronger correlation with instances' position on the map and is at most four for all our instances, indicating that they are all "close to stability".

*Summed Rank Minimal Stable Matchings.* We analyze summed rank minimal stable matchings, i.e., stable matchings $M$ minimizing $\sum_{a \in A} \mathrm{pos}_{\succ_a}(M(a))$ (these matchings are also sometimes called egalitarian stable matchings). Such matchings maximize the summed satisfaction of agents and are thus natural candidates to pick if multiple stable matchings exist. However, computing them is NP-hard [21] and thus we resorted to an ILP. For instances with a stable matching, we visualize the quality of summed rank minimal stable matchings in Figure 1e. Observe that instances sampled from one culture again behave remarkably similarly. In addition, there is some but certainly not a perfect correlation between the results and instances' position on the map: Ignoring Reverse-Euclidean instances which are a clear outlier here, if we move from chaos to mutual agreement the minimal summed rank decreases (as for perfect mutual agreement every agent can be matched to its top-choice); in contrast, if we move from chaos to mutual disagreement or from chaos to identity, then the minimal summed rank constantly increases.

*Further Types of Stable Matchings.* Complementing the results from the previous paragraph, we also examine the stable matching minimizing the summed satisfaction of agents. Interestingly, for almost all instances the lowest and highest possible summed satisfaction is similar, indicating that the space of stable matching is in some sense not very rich in these instances. We also study the stable matching maximizing the satisfaction of the agent worst off. Interestingly, here, models producing (close to) Euclidean instances show a diverse behavior, whereas for other models in all sampled instances the maximum satisfaction of the worst off agent is similar. Moreover, it is again possible to identify different regions on the corresponding maps (e.g., for instances close to identity, it is not possible to satisfy all agents adequately) as well as a continuously changing behavior when moving from some of the extreme points to others (e.g., moving from identity to mutual disagreement or chaos the situation of the worst-off agent constantly improves).

*Running Time Analysis.* Lastly, to illustrate another possible application of the map, in Figure 1f we visualize the time our ILP, which we solved using Gurobi Optimization, LLC [25], needed to find a summed rank minimal stable matching. Analyzing the results, again instances from the same culture behave quite similarly to each other. Moreover, the results are clearly connected to instances' position on the map. More specifically, instances from the Euclidean and Fame-Euclidean model seem to be particularly easy to solve, whereas instances close to ID and close to MD seem to be particularly challenging, maybe because here the achievable minimum summed rank is quite high. This is in contrast to election-related problems, where typically Impartial Culture elections are most challenging and the more structure there is in an election, the easier it is to solve [41].

## 6 A MAP OF STABLE MARRIAGE INSTANCES

So far we focused on the Stable Roommates problem. However, the developed framework can also be applied to other types of stable matching problems. To showcase this, we repeat parts of our studies for the Stable Marriage (SM) problem. Instances of SM differ from those of SR in that there is a bipartition of the agents into two size-$n$ sets (typically referred to as men and women) such that agents from one set have preferences over the agents from the other set and vice versa. In a matching, each pair then contains a man and a woman.

We extend the mutual attraction distance to SM instances as follows: Each instance now corresponds to two separate mutual attraction matrices $\mathcal{MA}^{\mathcal{I},U}$ and $\mathcal{MA}^{\mathcal{I},W}$, one ($\mathcal{MA}^{\mathcal{I},U}$) for men and one ($\mathcal{MA}^{\mathcal{I},W}$) for women (as for SR, each row corresponds to an agent and contains entries $\mathcal{MA}^{\mathcal{I}}(a,1),\dots,\mathcal{MA}^{\mathcal{I}}(a,n)$). When computing the distance between two SM instances $\mathcal{I}$ and $\mathcal{I}'$, we compute the summed distance between the matrices $\mathcal{MA}^{\mathcal{I},U}$ and $\mathcal{MA}^{\mathcal{I}',U}$ and the matrices $\mathcal{MA}^{\mathcal{I},W}$ and $\mathcal{MA}^{\mathcal{I}',W}$, and the summed distance between the matrices $\mathcal{MA}^{\mathcal{I},U}$ and $\mathcal{MA}^{\mathcal{I}',W}$ and the matrices $\mathcal{MA}^{\mathcal{I},W}$ and $\mathcal{MA}^{\mathcal{I}',U}$. The distance between $\mathcal{I}$ and $\mathcal{I}'$ is then the minimum of these two values. This in particular implies that we do not fix that "women" in one instance correspond to "women" in the other instances (as in one-to-one applications the two sides are often in some sense exchangeable), but instead map the two sides to each other such that the resulting distance is minimized.

To create a map of SM instances, we again sample 460 instances from canonical extensions of the statistical cultures used for SR to the SM setting. The resulting map of SM instances looks very similar to the one for SR instances depicted in Figure 1a. Moreover, we repeat the experiments that we conducted for SR, observing in most cases very similar results.[6] In particular, instances sampled from one model perform very similarly in most experiments, instances that are close according to the mutual attraction distance share similar characteristics, and the map groups close instances together, enabling us to identify regions on the map with distinct behaviors.

## 7 DISCUSSION

Contributing to the toolbox for experiments for stable matching problems, we have introduced the polynomial-time computable mutual attraction distance and analyzed its properties as well as the space it induces. As a second step, we have described a variety of statistical cultures for generating synthetic stable matching instances, which allow one to create diverse easily customizable test datasets. One specific application of these two contributions is our map of stable matching instances. We have verified that the produced map is meaningful in the sense that it groups instances with similar properties together, and have provided intuitive interpretations of the different regions on the map.

To demonstrate the capabilities of the map and our test dataset, we have conducted various exemplary experiments. Overall, our experimental results underline the importance of using diverse test data. Among others, we have observed that sampling preferences uniformly at random results in instances that behave very similarly (and often quite different than instances sampled from other models) and that such instances only cover a small part of the space of instances. Overall, this questions the common practice to only examine preferences sampled uniformly at random in an experimental analysis, as it is quite unclear whether the results of these experiments generalize.[7] Specifically, our results presented in Section 5 demonstrate the insufficiency of only using uniformly at random sampled preferences in SM and SR instances when analyzing properties of specific stable matchings and performances of algorithms. This is slightly worrisome given that in the past several papers have analyzed properties of different types of stable matchings [13, 14] and conducted performance evaluations of algorithms [14, 19, 23] in SM and SR instances only using random preferences.

Moreover, our experiments and analysis of the map also reveal that instances which have a small mutual attraction distance (and thus are close to each other on the map) tend to have similar properties. This underlines the usefulness of this distance measure to assess the similarity of instances. Furthermore, it highlights the capabilities of the map as a non-aggregate visualization tool: Instead of presenting experimental results by listing different (sometimes non-robust) statistical quantities, on the map, we can depict the results on an instance level, thereby showing the full picture. Using this, it is often possible to identify general high-level trends and typical behavior of instances from different parts of the space. In a similar vein, the map also supports the informed planning of more focused follow-up experiments, by looking for parts on the map that show an interesting behavior and analyzing the respective cultures in more detail. To use the map for these purposes, a meaningful placement of the instances on the map which groups similar instances together is vital. The maps shown in Figure 1 provide some first clear evidence that this is indeed the case, which also justifies the usage of the mutual attraction distance as a practically useful and sufficient distance measure.

[6] One main contrast concerns the difference between the maximum and minimum summed satisfaction of agents in a stable matching. For SR, this difference was usually very small. For some SM instances this difference is larger, indicating that the space of stable matchings for some of the sampled SM instances is "richer".

[7] Further, it is also unclear why instances with uniformly at random sampled preferences are particularly practically useful. Quite the contrary, there is some evidence that preferences in reality are often not drawn uniformly at random: In the general setting of agents ranking different alternatives, Boehmer et al. [6] and Boehmer and Schaar [10] analyzed real-world preference data from a variety of sources observing that only few of these match a random preference sampling.

## Acknowledgments

NB was supported by the DFG project ComSoc-MPMS (NI 369/22). KH was supported by the DFG project FPTinP (NI 369/16). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101002854). We are grateful to the anonymous *MATCH-UP* 2022 and *AAMAS* 2023 reviewers for their thoughtful, constructive, and helpful comments.

European Research Council
Established by the European Commission

## REFERENCES

[1] Ishan Agarwal and Richard Cole. 2022. Stable Matching: Choosing Which Proposals to Make. *CoRR* abs/2204.04162 (2022). arXiv:2204.04162 https://arxiv.org/abs/2204.04162

[2] Esther M. Arkin, Sang Won Bae, Alon Efrat, Kazuya Okamoto, Joseph S. B. Mitchell, and Valentin Polishchuk. 2009. Geometric stable roommates. *Inf. Process. Lett.* 109, 4 (2009), 219–224.

[3] Nayantara Bhatnagar, Sam Greenberg, and Dana Randall. 2008. Sampling stable marriages: why spouse-swapping won't work. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '08)*. SIAM, 1223–1232.

[4] Niclas Boehmer, Robert Bredereck, Edith Elkind, Piotr Faliszewski, and Stanislaw Szufa. 2022. Expected Frequency Matrices of Elections: Computation, Geometry, and Preference Learning. *CoRR* abs/2205.07831 (2022). https://doi.org/10.48550/arXiv.2205.07831 arXiv:2205.07831

[5] Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, and Rolf Niedermeier. 2021. Winner Robustness via Swap- and Shift-Bribery: Parameterized Counting Complexity and Experiments. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI '21)*. ijcai.org, 52–58.

[6] Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, Rolf Niedermeier, and Stanislaw Szufa. 2021. Putting a Compass on the Map of Elections. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI '21)*. ijcai.org, 59–65.

[7] Niclas Boehmer, Piotr Faliszewski, Rolf Niedermeier, Stanislaw Szufa, and Tomasz Was. 2022. Understanding Distance Measures Among Elections. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI '22)*. ijcai.org, 102–108.

[8] Niclas Boehmer, Klaus Heeger, and Rolf Niedermeier. 2022. Theory of and Experiments on Minimally Invasive Stability Preservation in Changing Two-Sided Matching Markets. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI '22)*. AAAI Press, 4851–4858.

[9] Niclas Boehmer, Klaus Heeger, and Stanislaw Szufa. 2022. A Map of Diverse Synthetic Stable Roommates Instances. *CoRR* abs/2208.04041 (2022). https://doi.org/10.48550/arXiv.2208.04041 arXiv:2208.04041

[10] Niclas Boehmer and Nathan Schaar. 2022. Collecting, Classifying, Analyzing, and Using Real-World Elections. *CoRR* abs/2204.03589 (2022). arXiv:2204.03589 https://arxiv.org/abs/2204.03589

[11] Robert Bredereck, Klaus Heeger, Dušan Knop, and Rolf Niedermeier. 2020. Multidimensional Stable Roommates with Master List. In *Proceedings of the 16th International Conferenc of the Web and Internet Economics (WINE '20)*. Springer, 59–73.

[12] Angelina Brilliantova and Hadi Hosseini. 2022. Fair Stable Matching Meets Correlated Preferences. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22)*. IFAAMAS, 190–198.

[13] Frances Cooper and David F. Manlove. 2019. Two-sided profile-based optimality in the stable marriage problem. *CoRR* abs/1905.06626 (2019). arXiv:1905.06626 http://arxiv.org/abs/1905.06626

[14] Frances Cooper and David F. Manlove. 2020. Algorithms for New Types of Fair Stable Matchings. In *Proceedings of the 18th International Symposium on Experimental Algorithms (SEA '20)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 20:1–20:13.

[15] Lin Cui and Weijia Jia. 2013. Cyclic stable matching for three-sided networking services. *Comput. Networks* 57, 1 (2013), 351–363.

[16] Maxence Delorme, Sergio García, Jacek Gondzio, Jörg Kalcsics, David F. Manlove, and William Pettersson. 2019. Mathematical models for stable matching problems with ties and incomplete lists. *Eur. J. Oper. Res.* 277, 2 (2019), 426–441.

[17] Persi Diaconis and Ronald L Graham. 1977. Spearman's footrule as a measure of disarray. *J. R. Stat. Soc. Series B Stat. Methodol.* 39, 2 (1977), 262–268.

[18] Cynthia Dwork, Ravi Kumar, Moni Naor, and D. Sivakumar. 2001. Rank aggregation methods for the Web. In *Proceedings of the Tenth International World Wide Web Conference (WWW '10)*. ACM, 613–622.

[19] Esra Erdem, Muge Fidan, David F. Manlove, and Patrick Prosser. 2020. A General Framework for Stable Roommates Problems using Answer Set Programming. *Theory Pract. Log. Program.* 20, 6 (2020), 911–925.

[20] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, Stanislaw Szufa, and Nimrod Talmon. 2019. How Similar Are Two Elections?. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI '19)*. AAAI Press, 1909–1916.

[21] Tomás Feder. 1992. A New Fixed Point Approach for Stable Networks and Stable Marriages. *J. Comput. Syst. Sci.* 45, 2 (1992), 233–284.

[22] D. Gale and L. S. Shapley. 2013. College Admissions and the Stability of Marriage. *Am. Math. Mon.* 120, 5 (2013), 386–391.

[23] Begum Genc, Mohamed Siala, Barry O'Sullivan, and Gilles Simonin. 2017. Finding Robust Solutions to Stable Marriage. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI '17)*. ijcai.org, 631–637.

[24] Begum Genc, Mohamed Siala, Gilles Simonin, and Barry O'Sullivan. 2019. An Approach to Robustness in the Stable Roommates Problem and Its Comparison with the Stable Marriage Problem. In *Proceedings of the 16th International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR 2019)*. Springer, 320–336.

[25] Gurobi Optimization, LLC. 2021. Gurobi Optimizer Reference Manual. https://www.gurobi.com. Accessed: 2022-04-25.

[26] Dan Gusfield and Robert W. Irving. 1989. *The Stable marriage problem - structure and algorithms.* MIT Press.

[27] Robert W. Irving and David F. Manlove. 2009. Finding large stable matchings. *ACM J. Exp. Algorithmics* 14 (2009), 1.2–1.30.

[28] Robert W. Irving, David F. Manlove, and Sandy Scott. 2008. The stable marriage problem with master preference lists. *Discret. Appl. Math.* 156, 15 (2008), 2959–2977.

[29] Tomihisa Kamada and Satoru Kawai. 1989. An Algorithm for Drawing General Undirected Graphs. *Inf. Process. Lett.* 31, 1 (1989), 7–15.

[30] Naoyuki Kamiyama. 2019. Many-to-Many Stable Matchings with Ties, Master Preference Lists, and Matroid Constraints. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '19)*. IFAAMAS, 583–591.

[31] Donald E. Knuth. 1976. *Mariages stables et leurs relations avec d'autres problèmes combinatoires.* Les Presses de l'Université de Montréal, Montreal, Que. 106 pages. Introduction à l'analyse mathématique des algorithmes, Collection de la Chaire Aisenstadt.

[32] Augustine Kwanashie and David F. Manlove. 2013. An Integer Programming Approach to the Hospitals/Residents Problem with Ties. In *Selected Papers of the International Conference on Operations Research (OR '13)*. Springer, 263–269.

[33] David F. Manlove. 2013. *Algorithmics of Matching Under Preferences.* Series on Theoretical Computer Science, Vol. 2. WorldScientific.

[34] David F. Manlove, Duncan Milne, and Sofiat Olaosebikan. 2022. Student-project allocation with preferences over projects: Algorithmic and experimental results. *Discret. Appl. Math.* 308 (2022), 220–234.

[35] Fredrik Manne, Md. Naim, Håkon Lerring, and Mahantesh Halappanavar. 2016. On Stable Marriages and Greedy Matchings. In *Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing (CSC '16)*. SIAM, 92–101.

[36] Stephan Mertens. 2015. Stable roommates problem with random preferences. *J. Stat. Mech. Theory Exp.* 2015, 1 (2015), P01020.

[37] William Pettersson, Maxence Delorme, Sergio García, Jacek Gondzio, Jörg Kalcsics, and David F. Manlove. 2021. Improving solution times for stable matching problems through preprocessing. *Comput. Oper. Res.* 128 (2021), 105128.

[38] Andrej Podhradsky. 2010. *Stable marriage problem algorithms.* Master's thesis. Masaryk University, Faculty of Informatics. is.muni.cz/th/172646/fi_m/

[39] Patrick Schober, Christa Boer, and Lothar A Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesth. Analg.* 126, 5 (2018), 1763–1768.

[40] Mohamed Siala and Barry O'Sullivan. 2017. Rotation-Based Formulation for Stable Matching. In *Proceedings of the 23rd International Conference on Principles and Practice of Constraint Programming (CP '17)*. Springer, 262–277.

[41] Stanislaw Szufa, Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. 2020. Drawing a Map of Elections in the Space of Statistical Cultures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '20)*. IFAAMAS, 1341–1349.

[42] Stanisław Szufa, Piotr Faliszewski, Łukasz Janeczko, Martin Lackner, Arkadii Slinko, Krzysztof Sornat, and Nimrod Talmon. 2022. How to Sample Approval Elections?. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI '22)*. ijcai.org, 496–502.

[43] Chung-Piaw Teo and Jay Sethuraman. 2000. On a cutting plane heuristic for the stable roommates problem and its applications. *Eur. J. Oper. Res.* 123, 1 (2000), 195–205.

[44] Nikolaos Tziavelis, Ioannis Giannakopoulos, Rune Quist Johansen, Katerina Doka, Nectarios Koziris, and Panagiotis Karras. 2020. Fair Procedures for Fair Stable Marriage Outcomes. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI '20)*. AAAI Press, 7269–7276.