

# Neutrality and Geometry of Mean Voting

Sébastien Lahaie, Microsoft Research, New York City, United States.

Nisarg Shah, Carnegie Mellon University, Pittsburgh, United States.

Mean proximity rules provide a simple geometric framework to achieve consensus among a collection of rankings (votes) over a set of alternatives. They embed all rankings into a Euclidean space, take the mean of the embeddings of the input votes, and return the ranking whose embedding is closest to the mean. Previous work on mean proximity rules has not integrated an important axiom—neutrality—into the framework. By drawing on ideas from the representation theory of finite groups, we show that integrating neutrality actually helps achieve a succinct representation for every mean proximity rule. Various connections are drawn between mean proximity rules and other prominent approaches to social choice.

Categories and Subject Descriptors: I.2.6 [Computing Methodologies]: Artificial Intelligence; J.4 [Computer Applications]: Social and Behavioral Sciences—*Economics*

Additional Key Words and Phrases: Social choice; Mean proximity rules; Euclidean space; Neutrality

## 1. INTRODUCTION

Modern social choice literature has seen rising interest in studying families of voting rules for aspects like manipulability [Gibbard 1973; Satterthwaite 1975; Xia 2013] and learnability [Procaccia et al. 2008; Procaccia et al. 2009]. Particularly, the framework of *generalized scoring rules* (GSRs) [Xia and Conitzer 2008] has been a great success in this direction. Two other well-studied approaches—*distance rationalizability* [Meskanen and Nurmi 2008; Elkind et al. 2009; 2010b] and *maximum likelihood estimation* (MLE) [de Condorcet 1785; Young 1988; Conitzer and Sandholm 2005]—have given rise to many families of voting rules as well. More recently, Caragiannis et al. [2013] introduced the families of PM-c and PD-c rules as generalizations of Condorcet consistent rules and positional scoring rules, respectively. Their goal was to design rules that perform well in predicting an underlying ground truth from noisy data.

These families of voting rules are extremely general in that generalized scoring rules, distance rationalizable rules, MLE rules, as well as the union of PM-c and PD-c rules contain almost all voting rules. Consequently, interesting restrictions of such families have been studied in the literature [Caragiannis et al. 2014; Elkind et al. 2009; 2010b; Conitzer and Sandholm 2005].

The family of *mean proximity rules* introduced by Zwicker [2008b] is one such restriction; it is a proper subset of each of generalized scoring rules, distance rationalizable rules, and maximum likelihood estimators (see Section 6). Additionally, each rule in the family does what a voting rule should intuitively do: aggregate the input votes into a mean consensus. A plethora of voting rules have emerged due to the fact that there is no natural notion of the *mean* of a set of rankings. Mean proximity rules embed rank-

---

Shah was visiting and partially supported by Microsoft Research New York City during the course of this work.

Author's addresses: S. Lahaie, Microsoft Research, New York City, slahaie@microsoft.com; N. Shah, Computer Science Department, Carnegie Mellon, nkshah@cs.cmu.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

EC'14, June 8–12, 2014, Stanford, CA, USA. Copyright © 2014 ACM 978-1-4503-2565-3/12/06 ...\$15.00.

<http://dx.doi.org/10.1145/2600057.2602898>

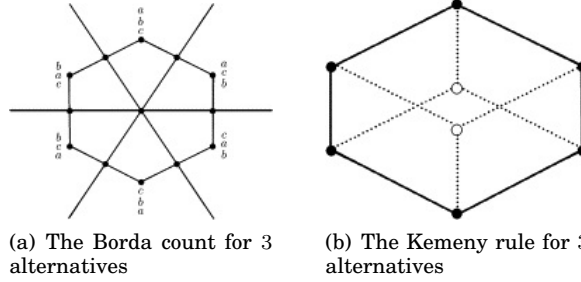


Fig. 1: Geometric visualizations of mean proximity rules.

ings into a Euclidean space, and use the standard mean (in Euclidean space) of the embeddings of the input votes to guide the consensus. When this mean does not represent embedding of any ranking, the ranking whose embedding is closest to the mean is chosen as the consensus. Different ways to embed the rankings into the Euclidean space give rise to different mean proximity rules. In this sense, mean proximity rules use Euclidean spaces to achieve the mean representative of the input votes. Examples of mean proximity rules include the Kemeny rule and all positional scoring rules.

Consider the embedding given in Figure 1(a) that generates the Borda count for 3 alternatives.<sup>1</sup> The 6 rankings are embedded into the vertices of a regular hexagon. Each voter adds a unit weight to the vertex corresponding to his vote, and the ranking whose region contains the mean (the center of gravity) of all the votes is chosen as the output. Similarly, Figure 1(b) shows the embedding generating the Kemeny rule for 3 alternatives. Mean proximity rules provide an easy understanding and in some cases simple visualizations of voting rules (even of the Kemeny rule that is usually considered a black-box algebraic problem). Saari [1995] argues that such a clear understanding is crucial for a voting rule to be accepted in the real-world, even more so for positional rules<sup>2</sup> that are arguably more complex than simple rules such as plurality or  $k$ -approval. He then makes a compelling case of using geometry in order to clarify and visualize voting rules. However, the geometric intuitions behind mean proximity rules are fundamentally different from those in Saari’s work.

Despite the attractive outlook, mean proximity rules have received less attention in the literature. We tap the potential and uncover significant structures in mean proximity rules. When introducing mean proximity rules, Zwicker [2008b] was motivated to remove neutrality primarily due to the characterization by Myerson [1995] that all social choice functions (voting rules that output a single winning alternative) satisfying neutrality, consistency, and another mild axiom are scoring rules. However, we study mean proximity rules in the context of social welfare functions (voting rules that output a ranking of alternatives), for which consistency is significantly weaker. Hence, neutrality can be integrated without much loss; for instance, the Kemeny rule is retained along with positional scoring rules. On the other hand, we show that integrating neutrality has substantial structural implications leading to benefits in terms of representation.

**Our contributions.** Mean proximity rules have two alternative representations, one geometric and one algebraic. We study mean proximity rules in a slightly limited context of social welfare functions, and characterize the constraints the limited context

<sup>1</sup>Image credit: [Ovchinnikov 2005].

<sup>2</sup>By positional rules, we mean rules that pay attention to alternatives in all positions of the input votes.

imposes on both representations (Theorem 4.9). Next, we study the important axiom of neutrality in the framework of mean proximity rules.

We give a characterization of the constraints imposed by neutrality on both representations of mean proximity rules (Theorem 5.9). In this context, we introduce *neutral embeddings*, and show that every neutral mean proximity rule admits a neutral embedding. By drawing ideas from the representation theory of finite groups, we further introduce *linear embeddings*, and show that every neutral mean proximity rule also admits a linear embedding. Linear embeddings are more constructive in nature, and thus have additional benefits over neutral embeddings. While storing a  $k$ -dimensional embedding for each of  $m!$  rankings typically takes  $O(k \cdot m!)$  space, linear embeddings leverage group theoretic structure to reduce the space requirement to  $O(k^2)$ . This provides a succinct representation for every neutral mean proximity rule, which distinguishes the family of low-dimensional neutral mean proximity rules from the other families of voting rules mentioned in the introduction where storing a rule requires exponential space in the worst case.

Using a fairly technical but constructive proof, in Theorem 5.15 we establish that (almost) every neutral embedding is also a linear embedding, and its succinct representation can also be computed. This leads to a new characterization of neutral mean proximity rules that yields further side results such as a characterization of mean proximity rules that admit low-dimensional embeddings. We conclude by drawing connections between mean proximity rules and various lines of research in social choice literature, showing that mean proximity rules connect well with each line of research.

## 2. RELATED WORK

We build upon the work of Zwicker [2008b; 2008a] who introduced and studied mean proximity rules and their generalization called mean neat rules, and gave an axiomatic characterization of (rational) mean neat voting rules. The work of Cervone et al. [2012] builds upon mean proximity rules and utilizes another notion of geometric consensus known as mediancenter or the Fermat-Weber point. They use the intuition of a rubber band and each voter exerting force to bring the consensus close to himself, which is similar to the intuition behind mean proximity rules.

The work of Saari [1995] uses a different geometric visualization where a profile is viewed in the space of  $m!$  dimensions, each dimension encoding the number of times a particular ranking appears in a given profile. Daugherty et al. [2009] extend his work by using intricate results from group representation theory. Related lines of research also include the work of Brams and Fishburn [2007] that argue in favor of approval voting for its simplicity and clarity, and the work of Crisman [2009] that connects the Borda count and the Kemeny rule using the geometry of permutahedron—the mean proximity embedding of the Borda count that is similar to Figure 1(a) but for 4 alternatives. Note that in these lines of research (as well as in this paper), the Euclidean space is merely used as a tool to understand voting rules, and does not bear any intrinsic meaning.

## 3. PRELIMINARIES

Let  $A$  be the set of alternatives, with cardinality  $|A| = m$ , and let  $n$  be the number of voters. Let  $\mathcal{L}(A)$  denote the set of rankings over the alternatives in  $A$ . A profile  $\pi \in \mathcal{L}(A)^n$  is a collection of votes (rankings). A *voting rule*, or more technically a *social welfare function* (SWF), is a function that maps every profile to a ranking or set of tied rankings.<sup>3</sup> Formally, we denote a voting rule by  $r : \mathcal{L}(A)^n \rightarrow \mathcal{P}(\mathcal{L}(A)) \setminus \{\emptyset\}$ , where  $\mathcal{P}(\cdot)$

<sup>3</sup>Another common definition of a voting rule is a *social choice function* (SCF), which maps every profile to a winning alternative or set of tied alternatives.

denotes the power set. Note that a voting rule must output at least one ranking so  $\emptyset$  is disallowed.

### Axiomatic Properties

Voting rules are often studied (and sometimes designed) from the viewpoint of axiomatic properties. Such properties prescribe the behavior of a voting rule in certain special cases. Some widely studied properties of relevance to our work are defined below.

*Anonymity.* A voting rule  $r$  is called *anonymous* if its output only depends on the collection of rankings in the profile (equivalently, the number of times each ranking appears), and not on the identities of the voters providing the different rankings.

*Consistency.* A voting rule  $r$  is called *consistent* if for profiles  $\pi_1$  and  $\pi_2$  such that  $r(\pi_1) \cap r(\pi_2) \neq \emptyset$ , we have  $r(\pi_1 + \pi_2) = r(\pi_1) \cap r(\pi_2)$ .<sup>4</sup> Here, profile  $\pi_1 + \pi_2$  denotes the union of the profiles  $\pi_1$  and  $\pi_2$ .

*Unanimity.* A voting rule  $r$  is said to satisfy *unanimity* if on every profile  $\pi$  that consists of copies of a single ranking  $\sigma$ , the rule uniquely outputs that ranking, so that  $r(\pi) = \{\sigma\}$ .

*Rank-distinguishability.* A voting rule  $r$  satisfies rank-distinguishability if for every two distinct rankings  $\sigma, \sigma' \in \mathcal{L}(A)$ , there exists a profile  $\pi$  such that exactly one of  $\sigma$  and  $\sigma'$  belongs to  $r(\pi)$ .

In the classical definition of social welfare functions that output a single ranking, non-imposition (also known as citizen sovereignty) of a voting rule dictates that every ranking of the alternatives should be achievable as the output of the voting rule on some profile. The rank-distinguishability axiom is a mild generalization of non-imposition to social welfare functions that output a set of rankings. It follows from unanimity which is itself considered unrestrictive and almost always desirable.

The next axiom will be central to our study. Let us choose the labels of the alternatives to be  $1, 2, \dots, m$ . Throughout we will denote the index set  $\{1, 2, \dots, m\}$  by  $[m]$ . A relabeling of the alternatives is equivalent to a permutation of  $[m]$ , namely an element of the symmetric group  $S_m$ . For a permutation  $\tau \in S_m$  and a ranking  $\sigma \in \mathcal{L}(A)$ , the ranking  $\tau\sigma$  is obtained by relabeling the alternatives according to  $\tau$  in  $\sigma$ . Given a profile  $\pi = (\sigma_1, \dots, \sigma_n)$  and a permutation  $\tau$  of the alternatives, let  $\tau\pi = (\tau\sigma_1, \dots, \tau\sigma_n)$  be the profile where each vote is permuted according to  $\tau$ . Similarly, given any set of rankings  $T$ , let  $\tau T = \{\tau\sigma : \sigma \in T\}$ .

*Neutrality.* A voting rule  $r$  is *neutral* if for every profile  $\pi$  and permutation  $\tau$  of the alternatives, we have  $r(\tau\pi) = \tau r(\pi)$ .

In words, a voting rule is called neutral if the labels of the alternatives do not matter, i.e., relabeling the alternatives would relabel the output of the rule (each ranking in the set returned) accordingly.

### Voting Rules

Next, we define two prominent voting rules that play a key role in this paper.

<sup>4</sup>Consistency for social welfare functions that output a set of tied rankings is more general than consistency for social welfare functions that output a single ranking (i.e., a singleton set). It is also significantly different from consistency of the winning alternative for social choice functions.

*Kemeny rule.* Given a profile  $\pi$ , define the weighted pairwise majority (PM) graph of  $\pi$  as the graph where the alternatives are the vertices and, for every pair of alternatives  $a$  and  $b$ , there is a directed edge from  $a$  to  $b$  with weight equal to the number of voters that prefer  $a$  to  $b$ . The Kemeny score of a ranking  $\sigma$  is the total weight of the edges in the PM graph of  $\pi$  whose direction agrees with  $\sigma$ . The Kemeny rule selects the ranking or set of tied rankings with the highest Kemeny score.

*Positional scoring rules.* A positional scoring rule is given by a scoring vector  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$  where  $\alpha_i \geq \alpha_{i+1}$  for all  $i \in [m]$  and  $\alpha_1 > \alpha_m$ . Under this rule, for each vote  $\sigma$  in  $\pi$  and  $i \in [m]$ ,  $\alpha_i$  points are awarded to the  $i$ th most preferred alternative in  $\sigma$ . The rule returns set of all rankings where the alternatives are sorted in descending order of total points. Some examples of positional scoring rules include plurality, Borda count, veto, and  $k$ -approval.

#### 4. MEAN PROXIMITY RULES

We begin by presenting important background on mean proximity rules that will be relevant for the rest of the paper. Mean proximity rules were originally defined in a scope greater than that of the voting rules considered in this paper: While voting rules map every profile to a set of tied rankings, mean proximity rules in general map every profile to a set of tied outcomes, where the outcomes are chosen from a generic outcome space  $\mathcal{O}$ . In Section 4.1, we will fix  $\mathcal{O} = \mathcal{L}(A)$  to analyze mean proximity rules in the context of voting rules. Here and throughout,  $\|\cdot\|$  refers to the Euclidean norm.

*Definition 4.1 (Mean Proximity Rules [Zwicker 2008b]).* A mean proximity rule is given by an input embedding  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  mapping input rankings to  $k$ -dimensional Euclidean space and an output embedding  $\psi : \mathcal{O} \rightarrow \mathbb{R}^k$  mapping possible outcomes to the same Euclidean space. The outcome of the rule on a profile  $\pi$  is given by the set  $\arg \min_{o \in \mathcal{O}} \|\psi(o) - \mu(\pi)\|$ , where  $\mu(\pi) = (1/n) \cdot \sum_{\sigma \in \pi} \phi(\sigma)$  is the mean of input embeddings of the votes in  $\pi$ , and  $\|\cdot\|$  is the Euclidean norm.<sup>5</sup>

Zwicker [2008b] also defined another family of rules, which he referred to as *generalized scoring rules*. However, this name has since been used primarily to represent a different family of voting rules that was introduced by Xia and Conitzer [2008]. We therefore refer to the family of Zwicker [2008b] as *matrix scoring rules*. In Section 6, we observe that the family of matrix scoring rules is a subset of the family of generalized scoring rules introduced by Xia and Conitzer [2008].

*Definition 4.2 (Matrix Scoring Rules [Zwicker 2008b]).* A voting rule is called a *matrix scoring rule* if there exists a scoring function  $s : \mathcal{L}(A) \times \mathcal{O} \rightarrow \mathbb{R}$  such that for any profile  $\pi$ , we have  $r(\pi) = \arg \max_{o \in \mathcal{O}} \sum_{\sigma \in \pi} s(\sigma, o)$ .

In words, the matrix scoring rule assigns a score to each outcome for each input vote, and the outcome or outcomes with the highest aggregate score are chosen. Our choice of terminology stems from the fact that the scoring function can be represented as a  $|\mathcal{L}| \times |\mathcal{O}|$  matrix with real entries. Zwicker [2008b] also showed equivalence between mean proximity rules and matrix scoring rules using the following technical result.

**PROPOSITION 4.3 ([ZWICKER 2008B]).** *For embeddings  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  and  $\psi : \mathcal{O} \rightarrow \mathbb{R}^k$ , and profile  $\pi$ , we have*

$$\arg \min_{o \in \mathcal{O}} \|\psi(o) - \mu(\pi)\| = \arg \min_{o \in \mathcal{O}} \sum_{\sigma \in \pi} \|\psi(o) - \phi(\sigma)\|^2. \quad (1)$$

<sup>5</sup>In this paper, every summation of the form  $\sum_{\sigma \in \pi}$  iterates over rankings in profile  $\pi$ ; in particular, if a ranking appears multiple times in  $\pi$ , the summation contains one term for each appearance.

From the definition of matrix scoring rules and Proposition 4.3, it is clear that the scoring function  $s(\sigma, o) = -\|\psi(o) - \phi(\sigma)\|^2$  represents the same mean proximity rule that has the embeddings  $(\psi, \phi)$ . Extending this to a two-way correspondence, Zwicker [2008b] showed the following.

**PROPOSITION 4.4** ([ZWICKER 2008B, THEOREM 4.2.1]). *The families of mean proximity rules and matrix scoring rules coincide.*

According to this result, a mean proximity rule has two equivalent representations: a pair of embeddings  $(\psi, \phi)$ , and a scoring function  $s$ ; it may be the case that neither of the two is unique. We conclude the background on mean proximity rules by mentioning a result of Zwicker [2008a] that further motivates the study of mean proximity rules.

**PROPOSITION 4.5** ([ZWICKER 2008A]). *All positional scoring rules and the Kemeny rule are mean proximity rules, with the outcome space being the set of rankings over the alternatives. A mean proximity rule is consistent, connected, continuous, and anonymous.*

Consistency of mean proximity rules is evident from Equation (1) because if two profiles lead to overlapping outputs according to the rule, then this overlap minimizes the sum on the right-hand side of Equation (1) for each profile, and therefore for their union. Connectedness and continuity are two natural properties of voting rules defined in [Zwicker 2008b]. For an alternative but related definition of continuity, the reader may refer to [Conitzer et al. 2009]. Conitzer and Sandholm [2005] showed that other rules such as Bucklin’s rule, Copeland’s method, the maximin rule, and the ranked pairs method are not consistent as social welfare functions. Hence, these rules are not mean proximity rules.

#### 4.1. Symmetric Mean Proximity Rules

We now begin our investigation of mean proximity rules in the context of social welfare functions, which output rankings over alternatives. Thus, from here onwards, we fix  $\mathcal{O} = \mathcal{L}(A)$ . This has the following effect on the various equivalent representations of mean proximity rules.

- (1) The embeddings  $\phi, \psi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  both map rankings to a Euclidean space.
- (2) The scoring function  $s : \mathcal{L}(A) \times \mathcal{L}(A) \rightarrow \mathbb{R}$  describes *similarity* between two rankings. This special case was also defined and studied independently by Conitzer et al. [2009], who referred to such rules as *simple ranking scoring functions* (SRSFs).
- (3) Under a fixed enumeration  $\sigma_1, \dots, \sigma_{m!}$  of  $\mathcal{L}(A)$ , we can represent a scoring function by an  $m! \times m!$  *score matrix*  $S$  such that  $S_{ij} = s(\sigma_i, \sigma_j)$  for  $i, j \in [m!]$ .

Mean proximity rules constitute a broad family of voting rules, some of which violate natural desiderata. For example, a mean proximity rule given by a scoring function  $s$  that satisfies  $s(\sigma, \sigma') > s(\sigma, \sigma)$  for distinct  $\sigma, \sigma' \in \mathcal{L}(A)$  would not output  $\sigma$  even on the profile where all the votes are  $\sigma$ , violating unanimity. To solve this problem, we propose a simple adjustment.

**Definition 4.6** (*Symmetric Mean Proximity Rule*). We say that a voting rule is a *symmetric mean proximity rule* (SMPR) if there exists a mean proximity representation of the rule with identical input and output embeddings:  $\psi = \phi$ .

Since the outcome space for SWFs is identical to the space of input votes (the set of rankings over the alternatives),  $\psi = \phi$  is a natural restriction. Note that not all embeddings representing an SMPR satisfy  $\psi = \phi$ . We call the embeddings that satisfy

this restriction *symmetric embeddings*. From here onwards, we always represent an SMPR using a symmetric embedding  $(\phi, \phi)$ , written simply  $\phi$ .

In what follows, we assume that the SMPRs additionally satisfy rank-distinguishability. For SMPRs, rank-distinguishability is equivalent to the intuitive restriction that  $\phi$  must map all rankings to different points of the Euclidean space. To see this, note that if  $\phi$  mapped all rankings to distinct points, then the rule would output  $\{\sigma\}$  on a profile with a single vote  $\sigma$ . Thus, such a rule would achieve rank-distinguishability. Conversely, if  $\phi$  mapped two distinct rankings  $\sigma$  and  $\sigma'$  to the same point in Euclidean space, then it is easy to see that on every profile either both rankings would belong to the output of the rule or neither.

**OBSERVATION 4.7.** *An SMPR satisfies rank-distinguishability if and only if every symmetric embedding  $\phi$  representing the rule maps all rankings to distinct points of the Euclidean space.*

Under this assumption, it is easy to check that in a profile  $\pi$  where all the votes are  $\sigma$ , we have  $\mu(\pi) = \phi(\sigma)$ . As a result, the output of every symmetric mean proximity rule on  $\pi$  would be the singleton set  $\{\sigma\}$ . Therefore, not only does symmetry impose intuitive structure on mean proximity rules, it also helps achieve natural desiderata such as unanimity. Furthermore, the embeddings constructed in [Zwicker 2008b] for positional scoring rules and the Kemeny rule in fact satisfy the restriction  $\psi = \phi$ , showing that these canonical mean proximity rules are symmetric.

#### 4.2. Algebraic Characterization of Symmetry

Recall that mean proximity rules are equivalent to matrix scoring rules (Proposition 4.4). In the previous section we imposed symmetry on the embeddings of the mean proximity rules. It is therefore natural to ask: *What restriction does symmetry place on other equivalent representations of mean proximity rules?* For an embedding  $\phi$ , define the scoring function  $s_\phi$  such that  $s_\phi(\sigma, \sigma') = -\|\phi(\sigma) - \phi(\sigma')\|^2$  for all  $\sigma, \sigma' \in \mathcal{L}(A)$ . Then Equation (1) implies that  $s_\phi$  and  $\phi$  are equivalent, meaning they both represent the same SMPR. Further, the score matrix generated by  $s_\phi$  has a well-known structure.

**Definition 4.8 (Euclidean Distance Matrix).** A  $q \times q$  matrix  $A = (a_{ij})$  is called a *Euclidean distance matrix (EDM)* if there exist  $v_1, \dots, v_q \in \mathbb{R}^p$  ( $p \leq q$ ) such that  $a_{ij} = \|v_i - v_j\|^2$  for all  $i, j \in [q]$ .

A matrix  $A$  is *conditionally positive semidefinite* if  $x^T A x \geq 0$  for all vectors  $x$  such that  $\mathbf{1}^T x = 0$ , where  $\mathbf{1} = (1, 1, \dots, 1)$ . It is known that a non-positive matrix with zero diagonal entries is a negated EDM if and only if it is conditionally positive semidefinite; see [Ikramov and Savel'eva 2000, Thm 3.10] and [Schoenberg 1938].

By construction, the score matrix of  $s_\phi$  is a negated EDM and represents the given SMPR. Conversely, given any score matrix that is a negated EDM, by definition we can find a symmetric embedding  $\phi$  such that the score matrix is generated by the scoring function  $s_\phi$ . Hence, the rule represented by the score matrix must be an SMPR. This yields the following characterization.

**THEOREM 4.9.** *A voting rule is a symmetric mean proximity rule if and only if it is a matrix scoring rule whose score matrix is a negated EDM, or equivalently a non-positive, zero-diagonal, conditionally positive semidefinite matrix.*

Note that while mean proximity SWFs coincide with the family of matrix scoring rules (i.e., simple ranking scoring functions) due to Conitzer et al. [2009], our focus on *symmetric* mean proximity rules gives a non-trivial algebraic condition on the score matrix

that translates to an intuitive geometric condition in the alternative mean proximity representation.

To summarize, we provided two motivations for adding symmetry to mean proximity rules: (1) taking identical input and output embeddings reflects the fact that the outcome space coincides with the space of input votes in the case of voting rules, and (2) symmetric mean proximity rules achieve the additional requirement of unanimity while still capturing all well-known mean proximity rules. We then characterized symmetric mean proximity rules in three equivalent representations: embeddings in Euclidean space, scoring functions, and score matrices. In the next section, we provide characterizations under the addition of another desirable property, *neutrality*. While neutrality is also a mild requirement (all voting rules of interest are neutral), we show that it adds significant structure to mean proximity rules.

## 5. NEUTRALITY AND SYMMETRIC MEAN PROXIMITY RULES

Recall that neutrality of a voting rule states that the labels of the alternatives do not matter: relabeling alternatives relabels the output in the same fashion. We remark that our analysis of mean proximity rules focuses on understanding the way voting rules work *before* they break ties. Therefore, mean proximity rules output a set of tied results (in our case, rankings over the alternatives), and do not explicitly break ties. Note that most rules do not remain neutral after lexicographic tie-breaking is imposed, but are neutral (as per the definition in this paper) before ties are broken. Further, uniformly random tie-breaking (where each tied result is returned with equal probability) can always be used to retain neutrality.

In Section 5.1, we study the restrictions imposed by neutrality on the various equivalent representations of *symmetric* mean proximity rules. We connect neutrality of SMPRs with a notion of neutrality in the embedding, a similar notion of neutrality in the scoring function [Conitzer et al. 2009], and positive semidefiniteness of the score matrix. In Section 5.2, we give a constructive characterization by drawing on ideas from group representation theory.

### 5.1. Characterizations of Neutrality

We begin by defining neutrality of scoring functions as in [Conitzer et al. 2009].

*Definition 5.1 (Neutral Scoring Function).* A scoring function  $s : \mathcal{L}(A) \times \mathcal{L}(A) \rightarrow \mathbb{R}$  is called *neutral* if  $s(\tau\sigma, \tau\sigma') = s(\sigma, \sigma')$  for all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$  and permutations  $\tau \in S_m$ . We say that a score matrix is neutral if its underlying scoring function is neutral.

In words, a scoring function is neutral if the similarity between two rankings given by the function does not change when both rankings are permuted in the same way. Conitzer et al. [2009] showed that any scoring function of a neutral mean proximity rule can be converted to an equivalent neutral scoring function.

**PROPOSITION 5.2** ([CONITZER ET AL. 2009]). *Let  $s$  be a scoring function representing a mean proximity rule  $r$  that is neutral. Then, an equivalent neutral scoring function  $s^{NT}$  representing the same rule  $r$  can be defined as follows. For all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$  and permutations  $\tau \in S_m$ , let*

$$s^{NT}(\sigma, \sigma') = \sum_{\tau \in S_m} s(\tau\sigma, \tau\sigma'). \quad (2)$$

Additionally, it is easy to check that by definition every neutral scoring function represents a neutral mean proximity rule. We immediately obtain the following.



PROPOSITION 5.3 ([CONITZER ET AL. 2009, LEMMA 2]). *A mean proximity rule is neutral if and only if it is represented by a neutral scoring function.*

This proposition characterizes the conditions on the scoring function (and thus on the score matrix) that neutrality imposes on mean proximity rules. However, we are interested in the addition of neutrality to *symmetric* mean proximity rules. Together with Theorem 4.9, Proposition 5.3 implies that a voting rule is a symmetric mean proximity rule if and only if it has a neutral score matrix and a score matrix that is a negated EDM. *Is there a score matrix satisfying both conditions? What conditions does neutrality impose on the embedding of an SMPR?* To answer these questions, we first introduce a notion of neutrality for an embedding.

**Definition 5.4 (Neutral Embedding).** We say that an embedding  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  is *neutral* if for all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$  and permutations  $\tau \in S_m$ , we have  $\|\phi(\sigma) - \phi(\sigma')\| = \|\phi(\tau\sigma) - \phi(\tau\sigma')\|$ .

It can be checked that if the embedding  $\phi$  is neutral, then its corresponding score function  $s_\phi$  is also neutral. Further, in analogy to scoring functions, an embedding can also be neutralized.

LEMMA 5.5. *Let  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  be an embedding of an SMPR  $r$ . Then,  $\phi^{NT}$  defined as follows is an equivalent neutral embedding of the same SMPR  $r$ :*

$$\phi^{NT}(\sigma) = [\phi(\tau_1\sigma)^T \phi(\tau_2\sigma)^T \dots \phi(\tau_{m!}\sigma)^T]^T, \quad (3)$$

where  $\tau_1, \dots, \tau_{m!}$  is a fixed enumeration of the permutations in  $S_m$ . Further, the neutralization of  $\phi$  and the neutralization of its scoring function  $s_\phi$  are connected via  $s_{\phi^{NT}} = s_\phi^{NT}$ , where the neutral scoring function  $s_\phi^{NT}$  defined in (2) represents  $r$ .

PROOF. First, we show that  $\phi^{NT}$  is neutral. For all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$  and permutations  $\tau \in S_m$ ,

$$\begin{aligned} \|\phi^{NT}(\tau\sigma) - \phi^{NT}(\tau\sigma')\|^2 &= \sum_{\tau' \in S_m} \|\phi(\tau'\tau\sigma) - \phi(\tau'\tau\sigma')\|^2 \\ &= \sum_{\tau' \in S_m} \|\phi(\tau'\sigma) - \phi(\tau'\sigma')\|^2 = \|\phi^{NT}(\sigma) - \phi^{NT}(\sigma')\|^2, \end{aligned}$$

where the second equality holds because for  $\tau \in S_m$ , we have  $\{\tau'\tau : \tau' \in S_m\} = S_m$ , which is a property of any group. Hence,  $\phi^{NT}$  is neutral. Next, for all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$ ,

$$\begin{aligned} s_{\phi^{NT}}(\sigma, \sigma') &= -\|\phi^{NT}(\sigma) - \phi^{NT}(\sigma')\|^2 = -\sum_{\tau \in S_m} \|\phi(\tau\sigma) - \phi(\tau\sigma')\|^2 \\ &= \sum_{\tau \in S_m} s_\phi(\tau\sigma, \tau\sigma') = s_\phi^{NT}(\sigma, \sigma'). \end{aligned}$$

Thus, we have  $s_{\phi^{NT}} = s_\phi^{NT}$ . The scoring function  $s_\phi$  is equivalent to the embedding  $\phi$ , and thus also represents the rule  $r$ . From Proposition 5.2, the scoring function  $s_\phi^{NT} = s_{\phi^{NT}}$  also represents  $r$ . Hence, the embedding  $\phi^{NT}$  also represents  $r$ .  $\square$

We remark that the neutralized embedding  $\phi^{NT}$  is not very satisfactory because it maps rankings to a Euclidean space with dimension  $m!$  times the dimension used by  $\phi$ . (Of course, as there are only  $m!$  rankings, the embeddings lie on a subspace of dimension at most  $m!$ ) However, it has significant structure. For example, in addition

to neutrality we also have  $\|\phi^{NT}(\sigma)\|^2 = \sum_{\tau \in S_m} \|\phi(\tau\sigma)\|^2 = \sum_{\sigma' \in \mathcal{L}(A)} \|\phi(\sigma')\|^2$ . Hence,  $\|\phi^{NT}(\sigma)\|$  is independent of  $\sigma$ . That is,  $\phi^{NT}$  is an *equal norm embedding*.

**Definition 5.6 (Equal Norm Embedding).** We say that an embedding  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  has *equal norm* if  $\|\phi(\sigma)\| = \|\phi(\sigma')\|$  for all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$ .

For equal norm embeddings, squares of Euclidean distances in Proposition 4.3 can be replaced by inner products as follows.

**LEMMA 5.7.** *For an equal norm embedding  $\phi$  of an SMPR  $r$ , the scoring function  $s$  given by  $s(\sigma, \sigma') = \langle \phi(\sigma), \phi(\sigma') \rangle$  for all rankings  $\sigma, \sigma' \in \mathcal{L}(A)$  represents  $r$ .*

**PROOF.** Let the equal norm embedding  $\phi$  have  $\|\phi(\sigma)\| = c$  for all  $\sigma \in \mathcal{L}(A)$ . From Proposition 4.3, we know that on a profile  $\pi$ ,

$$\begin{aligned} r(\pi) &= \arg \min_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} \|\phi(\sigma) - \phi(\sigma')\|^2 \\ &= \arg \min_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} (c^2 + c^2 - 2\langle \phi(\sigma), \phi(\sigma') \rangle) \\ &= \arg \min_{\sigma \in \mathcal{L}(A)} 2c^2 \cdot m! - 2 \sum_{\sigma' \in \pi} \langle \phi(\sigma), \phi(\sigma') \rangle = \arg \max_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} \langle \phi(\sigma), \phi(\sigma') \rangle. \end{aligned}$$

Hence, by definition  $s$  is a scoring function representing  $r$ .  $\square$

The score matrix generated by the inner product scoring function of Lemma 5.7 has a well-known structure.

**Definition 5.8 (Gram Matrix).** A  $q \times q$  matrix  $A = (a_{ij})$  is called a *Gram matrix* if there exist vectors  $v_1, \dots, v_q \in \mathbb{R}^p$  ( $p \leq q$ ) such that  $a_{ij} = \langle v_i, v_j \rangle$  for all  $i, j \in [q]$ .

Recall that a matrix  $A$  is *positive semidefinite* if  $x^T A x \geq 0$  for all vectors  $x$ . It is a standard fact that a matrix is a Gram matrix if and only if it is positive semidefinite. With this machinery at hand, we are ready to answer the questions we posed regarding conditions on various representations of an SMPR imposed by neutrality.

**THEOREM 5.9.** *For a mean proximity rule  $r$ , the following are equivalent.*

- (1)  $r$  is neutral and symmetric.
- (2) There exists a neutral symmetric embedding representing  $r$ .
- (3) There exists a non-positive score matrix representing  $r$  which is neutral, conditionally positive semidefinite, and has zero diagonal entries.
- (4) There exists a score matrix representing  $r$  which is neutral, positive semidefinite, and has equal diagonal entries.

**PROOF.** It is easy to show that the second and the third conditions imply the first condition. If  $r$  has a score matrix satisfying the third conditions, namely it is a negated EDM, then by Proposition 5.3 and Theorem 4.9,  $r$  is a neutral SMPR. Similarly, if  $r$  is a symmetric mean proximity rule with a neutral embedding  $\phi$ , then for a profile  $\pi$ , we

have

$$\begin{aligned}
r(\tau\pi) &= \arg \min_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} \|\phi(\sigma) - \phi(\tau\sigma')\|^2 \\
&= \tau \arg \min_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} \|\phi(\tau\sigma) - \phi(\tau\sigma')\|^2 \\
&= \tau \arg \min_{\sigma \in \mathcal{L}(A)} \sum_{\sigma' \in \pi} \|\phi(\sigma) - \phi(\sigma')\|^2 = \tau r(\pi),
\end{aligned}$$

where the first transition follows from Proposition 4.3, and the third transition follows from neutrality of  $\phi$ . Thus,  $r$  is a neutral SMPR.

Conversely, let  $r$  be a neutral SMPR and  $\phi$  be its embedding. We proved in Lemma 5.5 that  $s_\phi^{NT} = s_{\phi^{NT}} = s$  (say). Then,  $s = s_\phi^{NT}$  implies that the score matrix corresponding to the scoring function  $s$  is neutral. Further,  $s = s_{\phi^{NT}}$  implies that the score matrix is also a negated EDM. Hence, the score matrix corresponding to the scoring function  $s$  is both neutral and a negated EDM, implying the third condition. Further, the equivalent embedding  $\phi^{NT}$  is neutral and represents  $r$ , implying the second condition. Thus, we have shown that the first three conditions are equivalent.

For equivalence with the fourth condition, take a neutral SMPR  $r$  and its embedding  $\phi$ . We saw that the embedding  $\phi^{NT}$  has equal norm, and hence the score matrix  $S$  corresponding to its scoring function  $s$  given by  $s(\sigma, \sigma') = \langle \phi^{NT}(\sigma), \phi^{NT}(\sigma') \rangle$  is a Gram score matrix representing  $r$  (Lemma 5.7). Since  $S$  is a Gram matrix, it is also positive semidefinite. The equal norm property of  $\phi^{NT}$  implies that  $S$  has equal diagonal entries. Finally, neutrality of  $\phi^{NT}$  implies neutrality of  $S$ . Hence,  $S$  satisfies the requirements of the fourth condition.

Conversely, take a mean proximity rule  $r$  with a score matrix  $S$  that is neutral, positive semidefinite, and has equal diagonal entries. Then, we can find an embedding  $\phi$  such that  $S$  is its Gram matrix. Neutrality and equal diagonal entries of  $S$  imply neutrality and the equal norm property of  $\phi$ . The latter implies that  $\phi$  also represents  $r$  (Lemma 5.7). Hence,  $r$  must be a neutral SMPR.  $\square$

Theorem 5.9 translates neutrality of an SMPR to neutrality of its embedding. While it is straightforward to check whether a given embedding is neutral, generating neutral embeddings is a non-trivial task. This limits the constructiveness of the characterization. Next, we improve the characterization by showing that neutrality of an embedding is equivalent to a structure that provides a way to *generate* neutral embeddings.

## 5.2. Linear Embeddings

Group representation theory has played a important role in a diverse set of fields including coding theory, quantum mechanics, and crystallography (in chemistry) [Barut and Raczka 1986; Kovalev et al. 1993; MacWilliams and Sloane 1977]. A linear representation of a group essentially maps every group element to a linear transformation of a vector space, i.e., to an element of the general linear group on the vector space. When a basis is chosen for the vector space, the linear transformation can be represented via a matrix. This is known as matrix representation of a group.

In our case, we draw on the representation theory of the symmetric group, which has gained much attention and for which many interesting structures have been discovered [James et al. 1984]. We fix the standard Euclidean basis for simplicity. We now introduce *linear embeddings* using ideas from linear representations of the symmetric group.

*Definition 5.10 (Linear Embeddings).* We say that an embedding  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$  is *linear* if there exists a representation function  $R : S_m \rightarrow \mathbb{R}^{k \times k}$ , called the representation of  $\phi$ , mapping each permutation to a  $k \times k$  real matrix such that

- (1) the identity permutation  $\tau_e$  is mapped to the identity matrix:  $R(\tau_e) = I_k$ ,
- (2)  $R$  respects the multiplication operator of the symmetric group:  $R(\tau_1 \tau_2) = R(\tau_1)R(\tau_2)$  for all permutations  $\tau_1, \tau_2 \in S_m$  (where  $\tau_1 \tau_2$  represents the multiplication of  $\tau_1$  and  $\tau_2$  within  $S_m$ ),
- (3)  $R$  maps permutations to orthogonal matrices:  $R(\tau)^T = R(\tau)^{-1}$  for all permutations  $\tau \in S_m$ , and
- (4) permuting a ranking is equivalent to rotating—because  $R$  uses orthogonal matrices—its embedding appropriately:  $\phi(\tau\sigma) = R(\tau)\phi(\sigma)$  for all rankings  $\sigma \in \mathcal{L}(A)$  and permutations  $\tau \in S_m$ .

It can be easily seen that the first and the second conditions together imply that  $R(\tau^{-1}) = R(\tau)^{-1}$  for all permutations  $\tau \in S_m$ , where  $\tau^{-1}$  is the inverse of  $\tau$  in  $S_m$ . Further, the third condition is unrestrictive: A striking result from group representation theory states that for a representation  $R$  of any finite group (not just the symmetric group), there exists a matrix  $P$  such that the equivalent representation  $R'$  defined by  $R'(g) = PR(g)P^{-1}$  for each group element  $g$  uses orthogonal matrices; see, e.g., [Boerner et al. 1963, Thm 6.3]. That is, any representation can be converted to a representation that uses orthogonal matrices. In what follows, we use  $R_\tau$  instead of  $R(\tau)$  for notational convenience.

Before we dive into the equivalence between neutrality and linearity, we build its pillars by presenting a few results about linear embeddings.

**LEMMA 5.11.** *For a linear embedding  $\phi$ , rankings  $\sigma, \sigma' \in \mathcal{L}(A)$ , and permutation  $\tau \in S_m$ , we have*

$$\langle \phi(\tau\sigma), \phi(\tau\sigma') \rangle = \langle \phi(\sigma), \phi(\sigma') \rangle.$$

**PROOF.** Let  $R$  be the representation of  $\phi$ . Then, we have

$$\langle \phi(\tau\sigma), \phi(\tau\sigma') \rangle = \langle R_\tau \phi(\sigma), R_\tau \phi(\sigma') \rangle = \langle \phi(\sigma), R_\tau^T R_\tau \phi(\sigma') \rangle = \langle \phi(\sigma), \phi(\sigma') \rangle, \quad (4)$$

where, the last transition holds because  $R_\tau$  is an orthogonal matrix by definition.  $\square$

Taking  $\sigma = \sigma'$  in Lemma 5.11, we get the following corollary.

**COROLLARY 5.12.** *Every linear embedding has equal norm.*

While every linear embedding has equal norm, not every neutral embedding has equal norm. However, the next result shows that every neutral embedding has a similar equal distance property. The proof considers a profile where each ranking appears exactly once; a detailed proof is given in the full version of the paper available from the authors.

**LEMMA 5.13.** *Let  $\phi$  be a neutral embedding, and  $\phi_{avg} = (1/m!) \cdot \sum_{\sigma \in \mathcal{L}(A)} \phi(\sigma)$  denote the mean embedding of all rankings under  $\phi$ . Then,  $\|\phi(\sigma) - \phi_{avg}\|$  is a constant independent of the ranking  $\sigma$ .*

Given Lemma 5.13, we can easily derive an equal norm embedding given a neutral embedding: Given a neutral embedding  $\phi$ , we say that the embedding  $\hat{\phi}$  defined by  $\hat{\phi}(\sigma) = \phi(\sigma) - \phi_{avg}$  for all  $\sigma \in \mathcal{L}(A)$  is the *normalization* of  $\phi$ . Note that  $\hat{\phi}$  is simply a translation applied to  $\phi$ , which is distance-preserving. Hence,  $\hat{\phi}$  represents the

same neutral SMPR as  $\phi$  does. This normalization achieves the equal norm property (Lemma 5.13), and also preserves neutrality and linearity of the original embedding. The proof is straightforward and given in the full version of the paper.

**LEMMA 5.14.** *An embedding is neutral (resp. linear) if and only if its normalization is neutral (resp. linear).*

Lemma 5.14 shows that the normalization of a neutral embedding is also neutral, in addition to having the equal norm property. We now show that this normalization is in fact linear—this requires a much deeper proof.

**THEOREM 5.15.** *An embedding is neutral if and only if its normalization is linear.*

**PROOF.** We first show that a linear embedding is also a neutral embedding. Let  $\phi$  be a linear embedding. Then for all  $\sigma, \sigma' \in \mathcal{L}(A)$  and  $\tau \in S_m$ , we have

$$\begin{aligned} \|\phi(\tau\sigma) - \phi(\tau\sigma')\|^2 &= \|\phi(\tau\sigma)\|^2 + \|\phi(\tau\sigma')\|^2 - 2 \cdot \langle \phi(\tau\sigma), \phi(\tau\sigma') \rangle \\ &= \|\phi(\sigma)\|^2 + \|\phi(\sigma')\|^2 - 2 \cdot \langle \phi(\sigma), \phi(\sigma') \rangle = \|\phi(\sigma) - \phi(\sigma')\|^2, \end{aligned}$$

where the second transition holds due to Lemma 5.11 and its Corollary 5.12. Thus,  $\phi$  is also neutral.

Now, for our main result, let  $\phi$  be an embedding whose normalization  $\hat{\phi}$  is linear. Then,  $\hat{\phi}$  is also neutral as shown above. Due to Lemma 5.14, this implies that  $\phi$  is also neutral.

The converse is much trickier, therefore we provide the high-level overview of the proof. Take a neutral embedding  $\phi : \mathcal{L}(A) \rightarrow \mathbb{R}^k$ , and let  $r$  be the rule it represents. We want to show that  $\hat{\phi}$  is linear. By Corollary 5.12 and Lemma 5.14, we know that  $\hat{\phi}$  is a ( $k$ -dimensional) equal norm neutral embedding. We first construct an equivalent embedding  $\phi^f$  as follows. Let  $k' \leq k$  be the dimension of the affine subspace of  $\mathbb{R}^k$  spanned by the points  $\{\hat{\phi}(\sigma)\}_{\sigma \in \mathcal{L}(A)}$ . Then, we obtain  $\phi^f$  by taking a distance-preserving mapping to  $\mathbb{R}^{k'}$ . Next, we show that  $\phi^f$  is linear by explicitly constructing its representation using Moore-Penrose pseudoinverses. Finally, we use this to construct an explicit representation for  $\hat{\phi}$  itself.

Now we provide the formal proof. First, it is easy to check that  $\hat{\phi}_{avg} = 0$ . Hence, the affine space spanned by the embeddings under  $\hat{\phi}$  form a linear subspace of  $\mathbb{R}^k$  of dimension  $k'$ . Take an orthonormal basis  $v_1, \dots, v_{k'}$  of this linear subspace, write  $\hat{\phi}(\sigma)$  as a linear combination of the  $k'$  basis vectors, and let  $\phi^f(\sigma)$  encode the coefficients of the linear combination. Equivalently, if  $P$  is the  $k' \times k$  matrix with  $v_1, \dots, v_{k'}$  as its rows, then  $\phi^f(\sigma) = P\hat{\phi}(\sigma)$  for all  $\sigma \in \mathcal{L}(A)$ . It is now standard to show that the transformation  $P$  preserves all distances. This has two implications: First,  $\phi^f$  represents the same neutral SMPR  $r$  as  $\hat{\phi}$  and  $\phi$  do. Second,  $\phi^f$  is also an equal norm neutral embedding like  $\hat{\phi}$ . These hold because the definitions of mean proximity rules, equal norm property, and neutrality of embeddings only use distances between pairs of embeddings.

By construction,  $\phi^f$  is of *full dimension*, i.e., the points where the rankings are mapped span the whole of  $\mathbb{R}^{k'}$ .<sup>6</sup> Next, consider the matrix  $A = [\phi^f(\sigma_1), \dots, \phi^f(\sigma_{m!})]$ , the  $k' \times m!$  matrix whose columns are the embeddings under  $\phi^f$ . Here,  $\sigma_1, \dots, \sigma_{m!}$  is an enumeration of  $\mathcal{L}(A)$ . This implies that  $A$  must have full row rank  $k'$ . Fix a permutation  $\tau \in S_m$ , and consider  $B = [\phi^f(\tau\sigma_1), \dots, \phi^f(\tau\sigma_{m!})]$ . Then for all  $i, j \in [m!]$ ,

$$(B^T B)_{ij} = \langle \phi^f(\tau\sigma_i), \phi^f(\tau\sigma_j) \rangle = \langle \phi^f(\sigma_i), \phi^f(\sigma_j) \rangle = (A^T A)_{ij}$$

<sup>6</sup>While it is not relevant for the proof, we remark that the fact that the  $m!$  points span  $\mathbb{R}^{k'}$  implies  $k' \leq m!$ .

where the second transition follows because  $\phi^f$  is neutral. Thus,  $A^T A = B^T B$ .

We now show that there exists a matrix  $R_\tau$  such that  $\phi^f(\tau\sigma) = R_\tau \phi^f(\sigma)$  for all  $\sigma \in \mathcal{L}(A)$ . Combining the equations for all  $\sigma \in \mathcal{L}(A)$ , it is equivalent to show the existence of an  $R_\tau$  such that  $R_\tau A = B$ , i.e., existence of a solution  $X$  of the linear system  $A^T X = B^T$ . It is well-known that the system of equations  $Ax = b$  has a solution if and only if  $AA^+b = b$ , where  $A^+$  is the Moore-Penrose pseudoinverse of  $A$  (see, e.g., [Barata and Hussein 2012]). Trivially extending this to multiple systems of linear equations, we can see that the necessary and sufficient condition for existence of the required  $R_\tau$  is  $A^T(A^T)^+B^T = B^T$ , or equivalently,  $BA^+A = B$ . The last derivation uses the fact that  $(A^T)^+ = (A^+)^T$ . Now,

$$\begin{aligned} B &= BB^+B = B((B^T B)^+ B^T) B = B(B^T B)^+ (B^T B) \\ &= B(A^T A)^+ (A^T A) = B((A^T A)^+ A^T) A = BA^+A. \end{aligned}$$

Refer [Barata and Hussein 2012] for the identities  $X = XX^+X$  (used in the first transition) and  $X^+ = (X^T X)^+ X^T$  (used in the second and the fifth transitions) regarding Moore-Penrose pseudoinverses. The fourth transition holds because  $A^T A = B^T B$ . Hence, we have shown that for every permutation  $\tau \in S_m$ , there exists a matrix  $R_\tau$  such that  $\phi^f(\tau\sigma) = R_\tau \phi^f(\sigma)$  for all  $\sigma \in \mathcal{L}(A)$ . Further, one solution of  $Ax = b$  is  $x = A^+b$ . Extending this, we can see that one solution of  $A^T X = B^T$  is  $X = (A^T)^+ B^T$ . Hence,  $R_\tau = X^T = BA^+$ . We choose this particular solution for every  $\tau \in S_m$  and show that it satisfies the conditions of linearity.

Recall that the rank of the product of two matrices is at most the minimum of the rank of the two matrices, and  $R_\tau A = B$ . Also, both  $A$  and  $B$  are rank  $k'$  matrices. Hence,  $\text{rank}(R_\tau) \geq k'$ . However,  $R_\tau$  is a  $k' \times k'$  matrix. Hence, we conclude that  $R_\tau$  is invertible for every  $\tau \in S_m$ . We now show that  $R_{\tau_1 \tau_2} = R_{\tau_1} R_{\tau_2}$  for all  $\tau_1, \tau_2 \in S_m$ .

Fix  $\tau_1, \tau_2 \in S_m$ . For every  $\sigma \in \mathcal{L}(A)$ , we have  $\phi^f(\tau_1 \tau_2 \sigma) = R_{\tau_1 \tau_2} \phi^f(\sigma)$ . Also,  $\phi^f(\tau_1 \tau_2 \sigma) = R_{\tau_1} \phi^f(\tau_2 \sigma) = R_{\tau_1} R_{\tau_2} \phi^f(\sigma)$ . Thus, we have that  $R_{\tau_1 \tau_2} \phi^f(\sigma) = R_{\tau_1} R_{\tau_2} \phi^f(\sigma)$  for all  $\sigma \in \mathcal{L}(A)$ . It is easy to show that this implies  $R_{\tau_1 \tau_2} = R_{\tau_1} R_{\tau_2}$ . To see this, consider the submatrix of  $A$ —let us call it  $A_p$ —by taking  $k'$  linearly independent columns of  $A$ . Then,  $A_p$  is an invertible square matrix. Also,  $R_{\tau_1 \tau_2} A_p = R_{\tau_1} R_{\tau_2} A_p$ . Multiplying by  $A_p^{-1}$  on both sides, we get the desired result. To complete the proof, we need to show that  $R_\tau$  is an orthogonal matrix for all  $\tau \in S_m$ , i.e.,  $R_\tau^{-1} = R_\tau^T$ . Note that

$$R_\tau^T B = (A^+)^T B^T B = (A^+)^T A^T A = (AA^+)^T A = IA = A,$$

where the second transition holds because  $B^T B = A^T A$ , and the fourth transition holds because  $AA^+ = I$  for any full row-rank matrix  $A$  (see, e.g., [Barata and Hussein 2012]). Thus, we have that  $R_\tau^T B = A$ . Applying  $R_\tau$  on both sides, we get  $R_\tau R_\tau^T B = R_\tau A = B$ . Similarly to  $A_p$ , we can construct an invertible submatrix  $B_p$  of  $B$ , and get  $R_\tau R_\tau^T B_p = B_p$ , which, upon right-multiplication by  $B_p^{-1}$ , yields the required equation  $R_\tau^{-1} = R_\tau^T$ . Thus,  $\phi^f$  is a linear embedding. We now show that  $\hat{\phi}$  is also linear.

Consider the orthonormal basis  $v_1, \dots, v_{k'}$  that was used for constructing  $\phi^f$  from  $\hat{\phi}$ , and the corresponding matrix  $P$  with  $v_i$ 's as its rows. Complete this to a basis  $v_1, \dots, v_k$  of  $\mathbb{R}^k$ , and consider the matrix  $Q$  that has all the  $v_i$ 's as its rows. Then,  $Q\hat{\phi}(\sigma) = [\phi^f(\sigma)^T, 0, \dots, 0]^T$ , where the number of zeros is  $k - k'$ . For a permutation  $\tau \in S_m$ , construct the  $k \times k$  matrix

$$R'_\tau = Q^T \begin{bmatrix} R_\tau & 0 \\ 0 & I \end{bmatrix} Q.$$

Here,  $R_\tau$  is a  $k' \times k'$  matrix and  $I$  is the  $(k - k') \times (k - k')$  identity matrix. Note that

$$R'_\tau \hat{\phi}(\sigma) = Q^T \begin{bmatrix} R_\tau & 0 \\ 0 & I \end{bmatrix} Q \hat{\phi}(\sigma) = Q^T \begin{bmatrix} R_\tau & 0 \\ 0 & I \end{bmatrix} [\phi^f(\sigma) 0 \dots 0]^T = Q^T [\phi^f(\tau\sigma)^T, 0, \dots, 0]^T = \hat{\phi}(\tau\sigma).$$

Thus,  $R'_\tau \hat{\phi}(\sigma) = \hat{\phi}(\tau\sigma)$  for all  $\tau \in S_m$  and  $\sigma \in \mathcal{L}(A)$ . Further,

$$\begin{aligned} R'_{\tau_1} R'_{\tau_2} &= Q^T \begin{bmatrix} R_{\tau_1} & 0 \\ 0 & I \end{bmatrix} Q Q^T \begin{bmatrix} R_{\tau_2} & 0 \\ 0 & I \end{bmatrix} Q = Q^T \begin{bmatrix} R_{\tau_1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} R_{\tau_2} & 0 \\ 0 & I \end{bmatrix} Q \\ &= Q^T \begin{bmatrix} R_{\tau_1} R_{\tau_2} & 0 \\ 0 & I \end{bmatrix} Q = Q^T \begin{bmatrix} R_{\tau_1 \tau_2} & 0 \\ 0 & I \end{bmatrix} Q = R'_{\tau_1 \tau_2}. \end{aligned}$$

Here, the second transition follows since  $QQ^T$  is identity matrix as rows of  $Q$  form an orthonormal basis. Because  $R'_\tau$  is product of three invertible matrices,  $R'_\tau$  itself is invertible. Further,

$$(R'_\tau)^T = Q^T \begin{bmatrix} R_\tau & 0 \\ 0 & I \end{bmatrix}^T Q = Q^T \begin{bmatrix} R_\tau^T & 0 \\ 0 & I \end{bmatrix} Q = Q^T \begin{bmatrix} R_{\tau^{-1}} & 0 \\ 0 & I \end{bmatrix} Q = R'_{\tau^{-1}}.$$

Thus,  $R'_\tau$  is also an orthogonal matrix. This completes the proof that  $\hat{\phi}$  is a linear embedding and  $R'$  is its representation.  $\square$

Theorem 5.9 showed that an SMPR is neutral if and only if it has a symmetric embedding that is neutral. The latter occurs if and only if the rule also has a symmetric embedding that is the normalization of a neutral embedding. Together with Theorem 5.15, this provides another characterization of neutral SMPRs.

**THEOREM 5.16.** *A symmetric mean proximity rule is neutral if and only if there is a linear embedding representing it.*

Theorem 5.16 should be viewed as a fifth equivalent condition of Theorem 5.9. Next, we show that Theorem 5.16 could also be derived using a much simpler proof, but later we describe the advantages of our equivalence result between neutrality and linearity over the alternative proof.

**LEMMA 5.17.** *For an embedding  $\phi$  of a neutral SMPR  $r$ , its neutralization  $\phi^{NT}$  given in Equation (3) is a linear embedding.*

The proof of Lemma 5.17 appears in the full version of the paper. Theorem 5.9 and Lemma 5.17 together provide another proof of Theorem 5.16. However, our original proof that showed a stronger connection (equivalence) between normalization of neutral embeddings and linear embeddings bears certain advantages described below.

The proof of Theorem 5.15 is constructive, meaning that one can use it to explicitly find the representation of the normalization of any given neutral embedding of an SMPR. The multiplicative property  $R_{\tau_1 \tau_2} = R_{\tau_1} R_{\tau_2}$  implies that one only needs to store the representation matrices corresponding to a generating set of  $S_m$ , and can generate the rest of the matrices by appropriate multiplications and inverses.<sup>7</sup>

It is well-known that the symmetric group  $S_m$  has a generating set of size two, irrespective of the value of  $m$ , consisting of: 1) a cycle, which shifts the labels of the alternatives in a cyclic fashion, and 2) a transposition, which exchanges the labels of two adjacent alternatives. The matrices corresponding to these two elements, along with the value of  $\phi(\sigma)$  for a single  $\sigma \in \mathcal{L}(A)$ , is enough to generate  $\phi(\sigma)$  for all  $\sigma \in \mathcal{L}(A)$ . Hence, if the dimension of  $\phi$  is  $k$ , the corresponding SMPR can be expressed using only  $2k^2 + k$  parameters instead of the  $k \cdot m!$  parameters required to naively store a  $k$ -dimensional vector for each of  $m!$  embeddings. To our knowledge, this is the first representation that is not directly exponential in  $m$ .

This succinct representation has many advantages. First, it reduces the storage space required to encode a neutral SMPR. Second, we believe that this succinct representation could also inform the design of efficient algorithms based on mean proximity

<sup>7</sup>A generating set of a group is a subset of the group such that every element of the group can be expressed as the product (under the group operation) of finitely many elements of the subset and their inverses.

rules. Finally, one aim in studying families of voting rules is to learn the most effective voting rule from the family, for which fewer parameters implies more efficient learning.

Further, Lemma 5.17 gives a linear embedding that has dimension  $m!$  times the dimension of the input embedding, whereas Theorem 5.15 produces a linear embedding of the same dimension as the input embedding. The lower dimension further emphasizes the benefits of the succinct representation shown above. *Our results show that every rule in a family of low-dimensional neutral mean proximity rules admits a succinct representation.* This is in sharp contrast with other families of voting rules studied in the social choice literature (see Section 1) where storing a rule may typically require space that is exponential in the number of alternatives. Note that positional scoring rules and the Kemeny rule can be expressed using a polynomial number of dimensions—respectively,  $O(m)$  and  $O(m^2)$ .

This discussion inspires an interesting basic question: *Which neutral SMPRs admit low dimensional embeddings?* Low dimensional embeddings are important because they provide simple intuition behind the working of the rules they represent. As it turns out, the structure of linear embeddings allow us to characterize all neutral SMPRs which admit one dimensional and two dimensional embeddings. One dimensional embeddings partition the set of rankings into two—the sets of rankings obtained by applying all odd and even permutations to any single ranking.<sup>8</sup> The embedding then maps all rankings in one set to the same point, thus mapping all possible rankings to one of two distinct points. The set of two dimensional embeddings can be mapped to points on a circle as follows. In the matrices corresponding to the generating set of  $S_m$  of size 2, it can be shown that the matrix corresponding to the cycle must be a rotation by  $4\pi/m!$ , whereas the matrix corresponding to the transposition must be a reflection around a line passing through the origin at angle  $\theta$ , where  $\theta \in [0, 2\pi)$ .<sup>9</sup> Further, we are able to show that for 3 alternatives, imposing Pareto efficiency in the winner and in the loser to the set of two dimensional embeddings yields a characterization of positional scoring rules over 3 alternatives.<sup>10</sup> We omit the proofs due to space constraints and because they are straightforward.

Table I summarizes the characterizations presented in the paper. In the case of multiple conditions within a table cell, each condition is equivalent to the required restriction on the mean proximity rule. A prime result that is omitted from the table is that an embedding  $\phi$  is neutral if and only if its normalization  $\hat{\phi} = \phi - \phi_{avg}$  is linear.

## 6. CONNECTIONS TO OTHER APPROACHES

In this section, we connect mean proximity rules with various lines of research in the social choice literature.

**Generalized scoring rules.** Using Equation (1), it can easily be shown that mean proximity rules (in our context of social welfare functions) are special cases of generalized scoring rules [Xia and Conitzer 2008]. Further, similar to the equivalent geometric (embedding) and algebraic (scoring function) representations of mean proximity rules, GSRs also admit equivalent algebraic and geometric representations; see [Mossel et al. 2013]. Finally, a generalization of mean proximity rules known as mean neat voting rules [Zwicker 2008a] is characterized by anonymity, consistency, and two other technical conditions (connectedness and continuity). In contrast, generalized scoring rules

<sup>8</sup>The parity (oddness or evenness) of a permutation  $\sigma$  is the parity of the number of inversions in  $\sigma$ , i.e., the number of pairs  $x, y$  such that  $x < y$  but  $\sigma(x) > \sigma(y)$ .

<sup>9</sup>Orientation of the  $4\pi/m!$  rotation as well as the initial embedding can be fixed without loss of generality.

<sup>10</sup>When all voters rank the same alternative first or the same alternative last, that alternative should be first or last respectively in the output.



<b>A mean proximity rule is</b>	<b>iff <math>\exists</math> a score matrix <math>S</math> such that</b>	<b>iff <math>\exists</math> embeddings <math>(\psi, \phi)</math> such that</b>
symmetric	$S$ is negation of a Euclidean distance matrix	$\psi = \phi$
symmetric and neutral	i) $S$ is neutral and negation of a Euclidean distance matrix ii) $S$ is neutral, positive semidefinite, and has equal diagonal entries	i) $\psi = \phi$ is neutral ii) $\psi = \phi$ is linear

Table I: Summary of the characterization results.

are characterized by anonymity and finite local consistency (a milder version of consistency). This provides an alternate proof that mean proximity rules are a subset of generalized scoring rules.

**Distance rationalizability.** Meskanen and Nurmi [2008] introduced the distance rationalizability framework where a rule is defined by a consensus class of profiles on which the output is already defined and a distance function between profiles. Given a profile, the corresponding distance rationalizable rule finds the closest consensus class and returns the output for that class. Elkind et al. [2010a] studied additively votewise distances, where the distance between two profiles is (informally) the sum of distances between their rankings. Consider the distance rationalizable rules with additively votewise distance, (strong) unanimity consensus classes, and where the distance between two rankings is squared Euclidean distance between their embeddings. These rules coincide with symmetric mean proximity rules.

**Maximum likelihood.** Given an embedding  $\phi$ , consider the noise model where the probability of observing a noisy sample  $\sigma$  given that the ground truth is  $\sigma^*$  is given by

$$\Pr[\sigma|\sigma^*] = \frac{e^{-\|\phi(\sigma) - \phi(\sigma^*)\|^2}}{\sum_{\sigma' \in \mathcal{L}(A)} e^{-\|\phi(\sigma') - \phi(\sigma^*)\|^2}} \propto e^{-\|\phi(\sigma) - \phi(\sigma^*)\|^2}. \quad (5)$$

This is a family of Gaussian noise models, of which the Mallows model is a special case [Mallows 1957]. It can be shown that in case of neutral embeddings, the normalization in the denominator is independent of the ground truth  $\sigma^*$ , similarly to the Mallows model. Doignon et al. [2004] showed that the Mallows model admits efficient sampling. It would be interesting to analyze what noise models arising from Equation (5) also admit efficient sampling.

## 7. DISCUSSION

An important question that is not yet well understood is the relation between various representations of the same mean proximity rule. Note that a mean proximity rule essentially takes a convex combination of the embeddings of the input votes, and returns the rankings whose embeddings are closest to the convex combination, i.e., it returns the set of rankings whose embeddings' Voronoi regions contain the convex combination. We conjecture that using Cauchy's rigidity theorem (see [Aigner and Ziegler 2010]), it can be established that two embeddings represent the same (neutral) symmetric mean proximity rule if and only if they are similar, i.e., obtained from one another by orthogonal transformations, translations, and scaling. This *unique representation conjecture*, if proven, would significantly deepen our understanding of mean voting in the Euclidean geometry.

Additionally, it can be shown that such a result would also entail our equivalence between neutrality and linearity (Theorem 5.15), and allow one to calculate the minimum dimension required by any embedding representing a given symmetric mean proximity rule. For instance, the conjecture would imply that the minimum dimension required is  $m - 1$  for any positional scoring rule and  $\binom{m}{2}$  for the Kemeny rule. More importantly, strong connections with rigidity theory and the geometry of Voronoi diagrams could then be leveraged, for example, to generalize the Euclidean distances used in mean proximity rules to Bregman divergences (see [Boissonnat et al. 2010]). Moreover, rigidity theory and computational geometry have recently seen significant progress in tackling algorithmic and computational challenges (see, e.g., [Biedl et al. 2007]), and this could inform the design of efficient algorithms based on mean proximity rules.

## REFERENCES

- M. Aigner and G. M. Ziegler. 2010. Cauchy's rigidity theorem. In *Proofs from THE BOOK*. Springer, 81–84.
- J. C. A. Barata and M. S. Hussein. 2012. The Moore–Penrose Pseudoinverse: A Tutorial Review of the Theory. *Brazilian Journal of Physics* 42, 1-2 (2012), 146–165.
- A. O. Barut and R. Raczká. 1986. *Theory of group representations and applications*. Vol. 2. World Scientific.
- T. Biedl, A. Lubiw, and M. Spriggs. 2007. Cauchy's theorem and edge lengths of convex polyhedra. In *Proceedings of the 10th International Conference on Algorithms and Data Structures (WADS)*. 398–409.
- H. Boerner, P. G. Murphy, J. Werner, and P. Mayer-Kalkschmidt. 1963. *Representations of groups: with special consideration for the needs of modern physics*. North-Holland.
- J. D. Boissonnat, F. Nielsen, and R. Nock. 2010. Bregman voronoi diagrams. *Discrete & Computational Geometry* 44, 2 (2010), 281–307.
- S. J. Brams and P. C. Fishburn. 2007. *Approval voting*. Springer.
- I. Caragiannis, A. D. Procaccia, and N. Shah. 2013. When do noisy votes reveal the truth?. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*. 143–160.
- I. Caragiannis, A. D. Procaccia, and N. Shah. 2014. Modal Ranking: A Uniquely Robust Voting Rule. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*. Forthcoming.
- D. P. Cervone, R. Dai, D. Gnoutcheff, G. Lanterman, A. Mackenzie, A. Morse, N. Srivastava, and W. S. Zwicker. 2012. Voting with rubber bands, weights, and strings. *Mathematical Social Sciences* 64, 1 (2012), 11–27.
- V. Conitzer, M. Rognlie, and L. Xia. 2009. Preference Functions That Score Rankings and Maximum Likelihood Estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*. 109–115.
- V. Conitzer and T. Sandholm. 2005. Common Voting Rules as Maximum Likelihood Estimators. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*. 145–152.
- K. D. Crisman. 2009. The Borda Count, the Kemeny Rule, and the Permutahedron. Manuscript. (2009).
- Z. Daugherty, A. K. Eustis, G. Minton, and M. E. Orrison. 2009. Voting, the symmetric group, and representation theory. *The American Mathematical Monthly* (2009), 667–687.
- M. de Condorcet. 1785. *Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix*. Imprimerie Royal. (1785). Facsimile published in 1972 by Chelsea Publishing Company, New York.
- J. P. Doignon, A. Pekeč, and M. Regenwetter. 2004. The repeated insertion model for rankings: Missing link between two subset choice models. *Psychometrika* 69, 1 (2004), 33–54.
- E. Elkind, P. Faliszewski, and A. Slinko. 2009. On Distance Rationalizability of Some Voting Rules. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*. 108–117.
- E. Elkind, P. Faliszewski, and A. Slinko. 2010a. On the role of distances in defining voting rules. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 375–382.
- E. Elkind, P. Faliszewski, and A. M. Slinko. 2010b. Good Rationalizations of Voting Rules.. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*. 774–779.
- A. Gibbard. 1973. Manipulation of Voting Schemes. *Econometrica* 41 (1973), 587–602.
- K. D. Ikramov and N. V. Savel'eva. 2000. Conditionally definite matrices. *Journal of Mathematical Sciences* 98, 1 (2000), 1–50.

- G. D. James, A. Kerber, P. M. Cohn, and G. de B. Robinson. 1984. *The representation theory of the symmetric group*. Vol. 16. Cambridge University Press Cambridge.
- O. V. Kovalev, H. T. Stokes, and D. M. Hatch. 1993. *Representations of the Crystallographic Space Groups: Irreducible Representations, Induced Representations, and Corepresentations*. Gordon and Breach Science.
- F. J. MacWilliams and N. J. A. Sloane. 1977. *The theory of error-correcting codes*. Vol. 16. Elsevier.
- C. L. Mallows. 1957. Non-null ranking models. *Biometrika* 44 (1957), 114–130.
- T. Meskanen and H. Nurmi. 2008. Closeness counts in social choice. In *Power, Freedom, and Voting*, M. Braham and F. Steffen (Eds.). Springer-Verlag.
- E. Mossel, A. D. Procaccia, and M. Z. Rácz. 2013. A Smooth Transition from Powerlessness to Absolute Power. *Journal of Artificial Intelligence Research* 48 (2013), 923–951.
- R. B. Myerson. 1995. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare* 12, 1 (1995), 59–74.
- S. Ovchinnikov. 2005. Hyperplane arrangements in preference modeling. *Journal of Mathematical Psychology* 49, 6 (2005), 481–488.
- A. D. Procaccia, A. Zohar, Y. Peleg, and J. S. Rosenschein. 2009. The learnability of voting rules. *Artificial Intelligence* 173, 12–13 (2009), 1133–1149.
- A. D. Procaccia, A. Zohar, and J. S. Rosenschein. 2008. Automated Design of Scoring Rules by Learning from Examples. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 951–958.
- D. Saari. 1995. *Basic geometry of voting*. Vol. 12. Springer.
- M. Satterthwaite. 1975. Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory* 10 (1975), 187–217.
- I. J. Schoenberg. 1938. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* 44 (1938), 522–536.
- L. Xia. 2013. Generalized scoring rules: a framework that reconciles Borda and Condorcet. *SIGecom Exchanges* 12, 1 (2013), 42–48.
- L. Xia and V. Conitzer. 2008. Generalized scoring rules and the Frequency of Coalitional Manipulability. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC)*. 109–118.
- H. P. Young. 1988. Condorcet's theory of voting. *The American Political Science Review* 82, 4 (1988), 1231–1244.
- W. S. Zwicker. 2008a. A characterization of the rational mean neat voting rules. *Mathematical and Computer Modelling* 48, 9 (2008), 1374–1384.
- W. S. Zwicker. 2008b. Consistency without neutrality in voting rules: When is a vote an average? *Mathematical and Computer Modelling* 48, 9 (2008), 1357–1373.