
Controlling the distance to a Kemeny consensus without computing it

Yunlong Jiao

MINES ParisTech – CBIO, PSL Research University, Institut Curie, INSERM, U900, Paris, France

YUNLONG.JIAO@MINES-PARISTECH.FR

Anna Korba

LTCI UMR No. 5141 Telecom ParisTech/CNRS, Institut Mines-Telecom Paris, 75013, France

ANNA.KORBA@TELECOM-PARISTECH.FR

Eric Sibony

LTCI UMR No. 5141 Telecom ParisTech/CNRS, Institut Mines-Telecom Paris, 75013, France

ERIC.SIBONY@TELECOM-PARISTECH.FR

Abstract

Due to its numerous applications, rank aggregation has become a problem of major interest across many fields of the computer science literature. In the vast majority of situations, Kemeny consensus(es) are considered as the ideal solutions. It is however well known that their computation is NP-hard. Many contributions have thus established various results to apprehend this complexity. In this paper we introduce a practical method to predict, for a ranking and a dataset, how close this ranking is to the Kemeny consensus(es) of the dataset. A major strength of this method is its generality: it does not require any assumption on the dataset nor the ranking. Furthermore, it relies on a new geometric interpretation of Kemeny aggregation that we believe could lead to many other results.

Among the many ways to state the rank aggregation problem stands out Kemeny aggregation (Kemeny, 1959). Defined as the problem of minimizing a cost function over the symmetric group (see Section 2 for the definition), its solutions, called Kemeny consensus(es), have been shown to satisfy desirable properties from many points of view (see for instance Young & Levenglick, 1978).

Computing a Kemeny consensus is however NP-hard, even for only four rankings (see Bartholdi et al., 1989; Cohen et al., 1999; Dwork et al., 2001). This fact has motivated the scientific community to introduce many approximation procedures and to evaluate them on datasets (see Schalekamp & van Zuylen, 2009; Ali & Meilă, 2012, for examples of procedures and experiments). It has also triggered a tremendous amount of work to obtain theoretical guarantees on these procedures and more generally to tackle the complexity of Kemeny aggregation from various perspectives. Some contributions have proven bounds on the approximation cost of procedures (see Diaconis & Graham, 1977; Coppersmith et al., 2006; Van Zuylen & Williamson, 2007; Ailon et al., 2008; Freund & Williamson, 2015) while some have established recovery properties (see for instance Saari & Merlin, 2000; Procaccia et al., 2012). Some other contributions have shown that exact Kemeny aggregation is tractable if some quantity is known on the dataset (see for instance Betzler et al., 2008; 2009; Cornaz et al., 2013) or if the dataset satisfies some conditions (see for instance Brandt et al., 2015). At last, some contributions have established approximation bounds that can be computed on the dataset (see Davenport & Kalagnanam, 2004; Conitzer et al., 2006; Sibony, 2014).

In this paper we introduce a novel approach to apprehend the complexity of Kemeny aggregation. We consider the following question: *Given a dataset and a ranking, can we predict how close the ranking is to a Kemeny consensus without computing the latter?* We exhibit a tractable quantity that allows to give a positive answer to this question.

1. Introduction

Given a collection of rankings on a set of alternatives, how to aggregate them into one ranking? This rank aggregation problem has gained a major interest across many fields of the scientific literature. Starting from elections in social choice theory (see for instance Borda, 1781; Condorcet, 1785; Arrow, 1950; Xia, 2015), it has been applied to meta-search engines (see for instance Dwork et al., 2001; Renda & Straccia, 2003; Desarkar et al., 2016), competitions ranking (see for instance Davenport & Lovell, 2005; Deng et al., 2014), analysis of biological data (see for instance Kolde et al., 2012; Patel et al., 2013) or natural language processing (see for instance Li, 2014; Zamani et al., 2014) among others.

The main practical application of our results is a simple method to obtain such a guarantee for the outcome of an aggregation procedure on any dataset. A major strength of our approach is its generality: it applies to all aggregation procedures, for any dataset.

Our results are based on a certain geometric structure of Kemeny aggregation (see Section 3) that has been barely exploited in the literature yet but constitutes a powerful tool. We thus take efforts to explain it in details. We believe that it could lead to many other results on Kemeny aggregation.

The paper is structured as follows. Section 2 introduces the general notations and states the problem. The geometric structure is detailed in Section 3 and further studied in Section 5 while our main result is presented in Section 4. At last, numerical experiments are described in details in Section 6 to address the efficiency and usefulness of our method on real datasets.

2. Controlling the distance to a Kemeny consensus

Let $\llbracket n \rrbracket = \{1, \dots, n\}$ be a set of alternatives to be ranked. A full ranking $a_1 \succ \dots \succ a_n$ on $\llbracket n \rrbracket$ is seen as the permutation σ of $\llbracket n \rrbracket$ that maps an item to its rank: $\sigma(a_i) = i$ for $i \in \llbracket n \rrbracket$. The set of all permutations of $\llbracket n \rrbracket$ is called the symmetric group and denoted by \mathfrak{S}_n . Given a collection of N permutations $\mathcal{D}_N = (\sigma_1, \dots, \sigma_N) \in \mathfrak{S}_n^N$, Kemeny aggregation aims at solving

$$\min_{\sigma \in \mathfrak{S}_n} C_N(\sigma) := \sum_{t=1}^N d(\sigma, \sigma_t), \quad (1)$$

where d is the Kendall's tau distance defined for $\sigma, \sigma' \in \mathfrak{S}_n$ as the number of their pairwise disagreements: $d(\sigma, \sigma') = \sum_{1 \leq i < j \leq n} \mathbb{I}\{(\sigma(j) - \sigma(i))(\sigma'(j) - \sigma'(i)) < 0\}$. The function C_N denotes the cost, and a permutation σ^* solving (1) is called a Kemeny consensus. We denote by \mathcal{K}_N the set of Kemeny consensuses on the dataset \mathcal{D}_N .

Exact Kemeny aggregation is NP-hard: it cannot be solved efficiently with a general procedure. This does not mean however that nothing can be done. For example, it is clear that on a dataset where all permutations are equal to a $\sigma_0 \in \mathfrak{S}_n$, the Kemeny consensus is trivially given by σ_0 . Many contributions from the literature have thus focused on a particular approach to apprehend some part of the complexity of Kemeny aggregation. The examples given in the introduction divide in three main categories.

- **General guarantees for approximation procedures.** These results provide a bound on the cost of one voting rule, valid for any dataset (see Diaconis & Graham, 1977; Coppersmith et al., 2006; Van Zuylen

& Williamson, 2007; Ailon et al., 2008; Freund & Williamson, 2015).

- **Bounds on the approximation cost computed from the dataset.** These results provide a bound, either on the cost of a consensus or on the cost of the outcome of a specific voting rule, that depends on a quantity computed from the dataset (see Davenport & Kalagnanam, 2004; Conitzer et al., 2006; Sibony, 2014).
- **Conditions for the exact Kemeny aggregation to become tractable.** These results ensure the tractability of exact Kemeny aggregation if the dataset satisfies some condition or if some quantity is known from the dataset (see for instance Betzler et al., 2008; 2009; Cornaz et al., 2013; Brandt et al., 2015).

In this paper, we introduce a novel approach to apprehend the complexity of Kemeny aggregation to some extent. More specifically, we consider the following problem:

The Problem. *Let $\sigma \in \mathfrak{S}_n$ be a permutation, typically output by a computationally efficient aggregation procedure on \mathcal{D}_N . Can we use computationally tractable quantities to give an upper bound for the distance $d(\sigma, \sigma^*)$ between σ and a Kemeny consensus σ^* on \mathcal{D}_N ?*

The answer to this problem is positive as we will elaborate. It is well known that the Kendall's tau distance takes its values in $\{0, \dots, \binom{n}{2}\}$ (see for instance Stanley, 1986). Our main result, Theorem 1, thus naturally takes the form: given σ and \mathcal{D}_N , if the proposed condition is satisfied for some $k \in \{0, \dots, \binom{n}{2} - 1\}$, then $d(\sigma, \sigma^*) \leq k$ for all consensuses $\sigma^* \in \mathcal{K}_N$. Its application in practice is then straightforward (see Section 4 for an illustration).

A major strength of our method is its generality: it can be applied to any dataset \mathcal{D}_N , any permutation σ . This is because it exploits a powerful geometric framework for the analysis of Kemeny aggregation.

3. Geometric analysis of Kemeny aggregation

Because of its rich mathematical structure, Kemeny aggregation can be analyzed from many different point of views. While some contributions deal directly with the combinatorics of the symmetric group (for instance Diaconis & Graham, 1977; Blin et al., 2011), some work for instance on the pairwise comparison graph (for instance Coppersmith et al., 2006; Conitzer et al., 2006; Jiang et al., 2011), and others exploit the geometry of the Permutahedron (for instance Saari & Merlin, 2000). In this paper, we analyze it via the Kemeny embedding (see also Jiao & Vert, 2015).

Definition 1 (Kemeny embedding). The Kemeny embed-

ding is the mapping $\phi : \mathfrak{S}_n \rightarrow \mathbb{R}^{\binom{n}{2}}$ defined by

$$\phi : \sigma \mapsto \begin{pmatrix} \vdots \\ \text{sign}(\sigma(j) - \sigma(i)) \\ \vdots \end{pmatrix}_{1 \leq i < j \leq n},$$

where $\text{sign}(x) = 1$ if $x \geq 0$ and -1 otherwise.

The Kemeny embedding ϕ maps a permutation to a vector in $\mathbb{R}^{\binom{n}{2}}$ where each coordinate is indexed by an (unordered) pair $\{i, j\} \subset [n]$ (we choose $i < j$ by convention). Though this vector representation is equivalent to representing a permutation as a flow on the complete graph on $[n]$, it allows us to perform a geometric analysis of Kemeny aggregation in the Euclidean space $\mathbb{R}^{\binom{n}{2}}$. Denoting by $\langle \cdot, \cdot \rangle$ the canonical inner product and $\|\cdot\|$ the Euclidean norm, the starting point of our analysis is the following result, already proven in [Barthelemy & Monjardet \(1981\)](#).

Proposition 1 (Background results). *For all $\sigma, \sigma' \in \mathfrak{S}_n$,*

$$\|\phi(\sigma)\| = \sqrt{\frac{n(n-1)}{2}} \text{ and } \|\phi(\sigma) - \phi(\sigma')\|^2 = 4d(\sigma, \sigma'),$$

and for any dataset $\mathcal{D}_N = (\sigma_1, \dots, \sigma_N) \in \mathfrak{S}_n^N$, Kemeny aggregation (1) is equivalent to the minimization problem

$$\min_{\sigma \in \mathfrak{S}_n} C'_N(\sigma) := \|\phi(\sigma) - \phi(\mathcal{D}_N)\|^2, \quad (2)$$

where

$$\phi(\mathcal{D}_N) := \frac{1}{N} \sum_{t=1}^N \phi(\sigma_t). \quad (3)$$

Remark 1. Proposition 1 says that Kemeny rule is a “Mean Proximity Rule”, a family of voting rules introduced in [Zwicker \(2008\)](#) and further studied in [Lahaie & Shah \(2014\)](#). Our approach actually applies more generally to other voting rules from this class but we limit ourselves to Kemeny rule in the paper for sake of clarity.

Proposition 1 leads to the following geometric interpretation of Kemeny aggregation, illustrated by Figure 1. First, as $\|\phi(\sigma)\| = \sqrt{n(n-1)/2}$ for all $\sigma \in \mathfrak{S}_n$, the embeddings of all the permutations in \mathfrak{S}_n lie on the sphere \mathbb{S} of center 0 and radius $R := \sqrt{n(n-1)/2}$. Notice that $\|\phi(\sigma) - \phi(\sigma')\|^2 = 4d(\sigma, \sigma')$ for all $\sigma, \sigma' \in \mathfrak{S}_n$ implies that ϕ is injective, in other words that it maps two different permutations to two different points on the sphere. A dataset $\mathcal{D}_N = (\sigma_1, \dots, \sigma_N) \in \mathfrak{S}_n^N$ is thus mapped to a weighted point cloud on this sphere, where for any $\sigma \in \mathfrak{S}_n$, the weight of $\phi(\sigma)$ is the number of times σ appears in \mathcal{D}_N . The vector $\phi(\mathcal{D}_N)$, defined by Equation (3), is then equal to the barycenter of this weighted point cloud. We call it the *mean embedding* of \mathcal{D}_N . Now, the reformulation of Kemeny aggregation given by Equation (2) means that a

Kemeny consensus is a permutation σ^* whose embedding $\phi(\sigma^*)$ is closest to $\phi(\mathcal{D}_N)$, with respect to the Euclidean norm in $\mathbb{R}^{\binom{n}{2}}$.

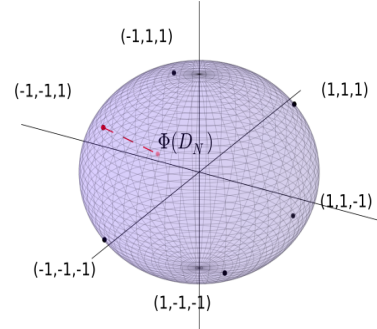


Figure 1. Kemeny aggregation for $n = 3$.

From an algorithmic point of view, Proposition 1 naturally decomposes problem (1) of Kemeny aggregation in two steps: first compute the mean embedding $\phi(\mathcal{D}_N)$ in the space $\mathbb{R}^{\binom{n}{2}}$, and then find a consensus σ^* as a solution of problem (2). The first step is naturally performed in $O(Nn^2)$ operations. The NP-hardness of Kemeny aggregation thus stems from the second step. In this regard, one may argue that having $\phi(\mathcal{D}_N)$ does not reduce much of the complexity in identifying an exact Kemeny consensus. However, a closer look at the problem leads us to asserting that $\phi(\mathcal{D}_N)$ greatly contains rich information about the localization of the Kemeny consensus(es). More specifically, we show in Theorem 1 that the knowledge of $\phi(\mathcal{D}_N)$ helps to provide an upper bound for the distance between a given permutation $\sigma \in \mathfrak{S}_n$ and any Kemeny consensus σ^* .

4. Main result

We now state our main result. For a permutation $\sigma \in \mathfrak{S}_n$, we define the angle $\theta_N(\sigma)$ between $\phi(\sigma)$ and $\phi(\mathcal{D}_N)$ by

$$\cos(\theta_N(\sigma)) = \frac{\langle \phi(\sigma), \phi(\mathcal{D}_N) \rangle}{\|\phi(\sigma)\| \|\phi(\mathcal{D}_N)\|}, \quad (4)$$

with $0 \leq \theta_N(\sigma) \leq \pi$ by convention.

Theorem 1. *Let $\mathcal{D}_N \in \mathfrak{S}_n^N$ be a dataset and $\sigma \in \mathfrak{S}_n$ a permutation. For any $k \in \{0, \dots, \binom{n}{2} - 1\}$, one has the following implication:*

$$\cos(\theta_N(\sigma)) > \sqrt{1 - \frac{k+1}{\binom{n}{2}}} \Rightarrow \max_{\sigma^* \in \mathcal{K}_N} d(\sigma, \sigma^*) \leq k.$$

The proof of Theorem 1 along with its geometric interpretation are postponed to Section 5. Here we focus on its application. Broadly speaking, Theorem 1 ensures that if the angle $\theta_N(\sigma)$ between the embedding $\phi(\sigma)$ of a permutation $\sigma \in \mathfrak{S}_n$ and the mean embedding $\phi(\mathcal{D}_N)$ is small, then the

Kemeny consensus(es) cannot be too far from σ . Its application in practice is straightforward. Assume that one applies an aggregation procedure on \mathcal{D}_N (say the Borda count for instance) with an output σ . A natural question is then: how far is it from the Kemeny consensus(es)? Of course, it is at most equal to $\max_{\sigma', \sigma'' \in \mathfrak{S}_n} d(\sigma', \sigma'') = \binom{n}{2}$. But if one computes the quantity $\cos(\theta_N(\sigma))$, it can happen that Theorem 1 allows to give a better bound. More specifically, the best bound is given by the minimal $k \in \{0, \dots, \binom{n}{2} - 1\}$ such that $\cos(\theta_N(\sigma)) > \sqrt{1 - (k+1)/\binom{n}{2}}$. Denoting by $k_{\min}(\sigma; \mathcal{D}_N)$ this integer, it is easy to see that

$$k_{\min}(\sigma; \mathcal{D}_N) = \begin{cases} \lfloor \binom{n}{2} \sin^2(\theta_N(\sigma)) \rfloor & \text{if } 0 \leq \theta_N(\sigma) \leq \frac{\pi}{2} \\ \lfloor \binom{n}{2} \cos^2(\theta_N(\sigma)) \rfloor & \text{if } \frac{\pi}{2} \leq \theta_N(\sigma) \leq \pi. \end{cases} \quad (5)$$

where $\lfloor x \rfloor$ denotes the integer part of the real x . We formalize this method in the following description.

Method 1. Let $\mathcal{D}_N \in \mathfrak{S}_n^N$ be a dataset and let $\sigma \in \mathfrak{S}_n$ be a permutation considered as an approximation of Kemeny's rule. In practice σ is the consensus returned by a tractable voting rule.

1. Compute $k_{\min}(\sigma; \mathcal{D}_N)$ with Formula (5).
2. Then by Theorem 1, $d(\sigma, \sigma^*) \leq k_{\min}(\sigma; \mathcal{D}_N)$ for any Kemeny consensus $\sigma^* \in \mathcal{K}_N$.

The following proposition ensures that Method 1 has tractable complexity.

Proposition 2 (Complexity of the method). *The application of Method 1 has complexity in time $O(Nn^2)$.*

With a concrete example, we demonstrate the applicability and the generality of Method 1.

Example 1 (Application of Method 1 on the sushi dataset.). We report here the results of a case-study on the sushi dataset provided by Kamishima (2003) to illustrate our method. The dataset consists of $N = 5000$ full rankings given by different individuals of the preference order on $n = 10$ sushi dishes such that a brute-force search for the Kemeny consensus is already quite computationally intensive. To apply our method, we select seven tractable voting rules, denoted by σ , as approximate candidates to Kemeny's rule to provide an initial guess (details of voting rules can be found in the Supplementary Material). Table 1 summarizes the values of $\cos(\theta_N(\sigma))$ and $k_{\min}(\sigma)$, respectively given by Equations (4) and (5). Results show that on this particular dataset, if we use for instance Borda Count to approximate Kemeny consensus, we are confident that the exact consensus(es) have a distance of at most 14 to the approximate ranking. We leave detailed interpretation of the results to Section 6.

Table 1. Summary of a case-study on the validity of Method 1 with the sushi dataset ($N = 5000, n = 10$). Rows are ordered by increasing k_{\min} (or decreasing cosine) value.

Voting rule	$\cos(\theta_N(\sigma))$	$k_{\min}(\sigma)$
Borda	0.820	14
Copeland	0.822	14
QuickSort	0.822	14
Plackett-Luce	0.80	15
2-approval	0.745	20
1-approval	0.710	22
Pick-a-Perm	0.383 [†]	34.85 [†]
Pick-a-Random	0.377 [†]	35.09 [†]

[†]For randomized methods such as Pick-a-Perm and Pick-a-Random, results are averaged over 10 000 computations.

5. Geometric interpretation and proof of Theorem 1

This section details the proof of Theorem 1 and its geometric interpretation. We deem that our proof has indeed a standalone interest, and that it could lead to other profound results on Kemeny aggregation.

5.1. Extended cost function

We recall that the Kemeny consensus(es) of a dataset \mathcal{D}_N are the solutions of Problem (2):

$$\min_{\sigma \in \mathfrak{S}_n} C'_N(\sigma) = \|\phi(\sigma) - \phi(\mathcal{D}_N)\|^2.$$

This is an optimization problem on the discrete set \mathfrak{S}_n , naturally hard to analyze. In particular the shape of the cost function C'_N is not easy to understand. However, since all the vectors $\phi(\sigma)$ for $\sigma \in \mathfrak{S}_n$ lie on the sphere $\mathbb{S} = \{x \in \mathbb{R}^{\binom{n}{2}} \mid \|x\| = R\}$ with $R = \sqrt{n(n-1)/2}$, it is natural to consider the relaxed problem on \mathbb{S}

$$\min_{x \in \mathbb{S}} C_N(x) := \|x - \phi(\mathcal{D}_N)\|^2. \quad (6)$$

We call C_N the extended cost function with domain \mathbb{S} . The advantage of C_N is that it has a very simple shape. We denote by $\theta_N(x)$ the angle between a vector $x \in \mathbb{S}$ and $\phi(\mathcal{D}_N)$ (with the slight abuse of notations that $\theta_N(\phi(\sigma)) \equiv \theta_N(\sigma)$). For any $x \in \mathbb{S}$, one has

$$C_N(x) = R^2 + \|\phi(\mathcal{D}_N)\|^2 - 2R\|\phi(\mathcal{D}_N)\| \cos(\theta_N(x)).$$

This means that the extended cost $C_N(x)$ of a vector $x \in \mathbb{S}$ only depends on the angle $\theta_N(x)$. The level sets of C_N are thus of the form $\{x \in \mathbb{S} \mid \theta_N(x) = \alpha\}$, for $0 \leq \alpha \leq \pi$. If $n = 3$, these level sets are circles in planes orthogonal to $\phi(\mathcal{D}_N)$, each centered around the projection of the latter on the plane (Figure 2). This property implies the following result.

Lemma 1. A Kemeny consensus of a dataset \mathcal{D}_N is a permutation σ^* such that:

$$\theta_N(\sigma^*) \leq \theta_N(\sigma) \quad \text{for all } \sigma \in \mathfrak{S}_n.$$

Lemma 1 means that the problem of Kemeny aggregation translates into finding permutations σ^* that have minimal angle $\theta_N(\sigma^*)$. This reformulation is crucial to our approach.

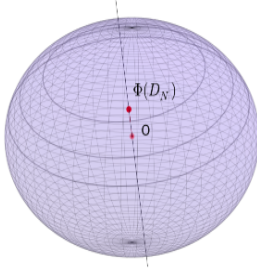


Figure 2. Level sets of \mathcal{C}_N over \mathbb{S} .

5.2. Interpretation of the condition in Theorem 1

The second element of our approach is motivated by the following observation. Let $x \in \mathbb{S}$ be a point on the sphere and let $r \geq 0$. If r is large enough, then all the points $x' \in \mathbb{S}$ on the sphere that have distance $\|x' - x\|$ greater than r will have a greater angle $\theta_N(x')$. Formally, we denote by $\mathcal{B}(x, r) = \{x' \in \mathbb{R}^{\binom{n}{2}} \mid \|x' - x\| < r\}$ the (open) ball of center x and radius r . Then one has the following result.

Lemma 2. For $x \in \mathbb{S}$ and $r \geq 0$, one has the following implication:

$$\cos(\theta_N(x)) > \sqrt{1 - \frac{r^2}{4R^2}} \Rightarrow \min_{x' \in \mathbb{S} \setminus \mathcal{B}(x, r)} \theta_N(x') > \theta_N(x).$$

Proof. Let $\bar{\phi}(\mathcal{D}_N) = \frac{\phi(\mathcal{D}_N)}{\|\phi(\mathcal{D}_N)\|}$. We discuss over two cases.

Case I: $\|\bar{\phi}(\mathcal{D}_N) - x\| \geq r$. By laws of cosines, this case is equivalent to:

$$\begin{aligned} 2R^2(1 - \cos(\theta_N(x))) &= \|\bar{\phi}(\mathcal{D}_N) - x\|^2 \geq r^2 \\ \Leftrightarrow \cos(\theta_N(x)) &\leq 1 - \frac{r^2}{2R^2} \leq 1 - \frac{r^2}{4R^2}. \end{aligned}$$

Note also that in this case, we have $\bar{\phi}(\mathcal{D}_N) \in \mathbb{S} \setminus \mathcal{B}(x, r)$ and hence $\min_{x' \in \mathbb{S} \setminus \mathcal{B}(x, r)} \theta_N(x') = \min_{x' \in \mathbb{S}} \theta_N(x') = 0 \leq \theta_N(x)$ always holds, where the minimum is attained at $x' = \bar{\phi}(\mathcal{D}_N)$.

Case II: $\|\bar{\phi}(\mathcal{D}_N) - x\| < r$, that is $\bar{\phi}(\mathcal{D}_N) \in \mathcal{B}(x, r)$. As the function $x' \mapsto \theta_N(x')$ is convex with global minimum in $\mathcal{B}(x, r)$, its minimum over $\mathbb{S} \setminus \mathcal{B}(x, r)$ is attained

at the boundary $\mathbb{S} \cap \partial\mathcal{B}(x, r) = \{x' \in \mathbb{R}^{\binom{n}{2}} \mid \|x'\| = R \text{ and } \|x' - x\| = r\}$, which is formed by cutting \mathbb{S} with the $((\binom{n}{2}) - 1)$ -dimensional hyperplane written as

$$\mathbb{L} := \left\{x' \in \mathbb{R}^{\binom{n}{2}} \mid \langle x', x \rangle = \frac{2R^2 - r^2}{2}\right\}$$

Straightforwardly one can verify that $\mathbb{S} \cap \partial\mathcal{B}(x, r)$ is in fact a $((\binom{n}{2}) - 1)$ -dimensional sphere lying in \mathbb{L} , centered at $c = \frac{2R^2 - r^2}{2R^2}x$ with radius $\gamma = r\sqrt{1 - \frac{r^2}{4R^2}}$. Now we take effort to identify:

$$x^* = \operatorname{argmin}_{x' \in \mathbb{S} \cap \partial\mathcal{B}(x, r)} \theta_N(x') = \operatorname{argmin}_{x' \in \mathbb{S} \cap \partial\mathcal{B}(x, r)} \mathcal{C}_N(x').$$

Note that $\phi(\mathcal{D}_N)$ projected onto \mathbb{L} is the vector $(\phi(\mathcal{D}_N))_{\mathbb{L}} := \phi(\mathcal{D}_N) - \frac{\langle \phi(\mathcal{D}_N), x \rangle}{R^2}x$. One can easily verify by Pythagoras rule that, for any set $\mathbb{K} \subseteq \mathbb{L}$,

$$\operatorname{argmin}_{x' \in \mathbb{K}} \|x' - \phi(\mathcal{D}_N)\| = \operatorname{argmin}_{x' \in \mathbb{K}} \|x' - (\phi(\mathcal{D}_N))_{\mathbb{L}}\|.$$

Therefore we have:

$$\begin{aligned} x^* &= \operatorname{argmin}_{x' \in \mathbb{S} \cap \partial\mathcal{B}(x, r)} \|x' - (\phi(\mathcal{D}_N))_{\mathbb{L}}\| = c + \gamma \frac{(\phi(\mathcal{D}_N))_{\mathbb{L}}}{\|(\phi(\mathcal{D}_N))_{\mathbb{L}}\|} \\ &= \frac{2R^2 - r^2}{2R^2}x + r\sqrt{1 - \frac{r^2}{4R^2}} \frac{\phi(\mathcal{D}_N) - \frac{\langle \phi(\mathcal{D}_N), x \rangle}{R^2}x}{\sqrt{\|\phi(\mathcal{D}_N)\|^2 - \frac{\langle \phi(\mathcal{D}_N), x \rangle^2}{R^2}}}. \end{aligned}$$

Tedious but essentially undemanding calculation leads to

$$\begin{aligned} \theta_N(x^*) > \theta_N(x) &\Leftrightarrow \langle x^*, \phi(\mathcal{D}_N) \rangle > \langle x, \phi(\mathcal{D}_N) \rangle \\ &\Leftrightarrow \cos(\theta_N(x)) > \sqrt{1 - \frac{r^2}{4R^2}}. \end{aligned}$$

□

It is interesting to look at the geometric interpretation of Lemma 2. In fact, it is clear from the proof that x^* should lie in the 2-dimensional subspace spanned by $\phi(\mathcal{D}_N)$ and x . We are thus able to properly define multiples of an angle by summation of angles on such linear space $2\theta_N(x) := \theta_N(x) + \theta_N(x)$. Figure 3 provides an illustration of Lemma 2 in this 2-dimensional subspace from the geometric point of view. In words, provided that $\theta_N(x) \leq \pi/2$, x^* has a smaller angle than x is equivalently written using laws of cosines as

$$\begin{aligned} r^2 = \|x - x^*\|^2 &> 2R^2(1 - \cos(2\theta_N(x))) \\ \Leftrightarrow \cos(2\theta_N(x)) &> 1 - \frac{r^2}{2R^2} \Leftrightarrow \cos(\theta_N(x)) > \sqrt{1 - \frac{r^2}{4R^2}}. \end{aligned}$$

This recovers exactly the condition stated in Lemma 2.

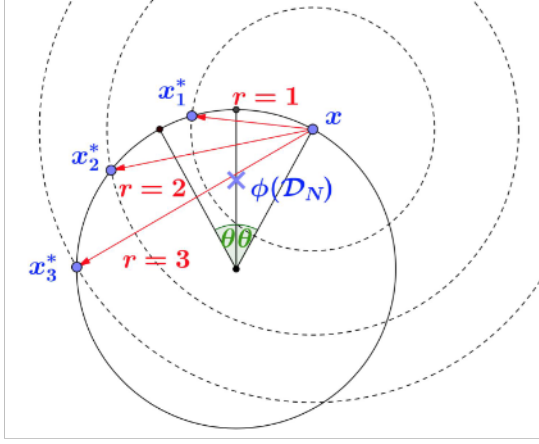


Figure 3. Illustration of Lemma 2 with r taking integer values (representing possible Kendall’s tau distance). The smallest integer value for r such that these inequalities hold is $r = 2$.

5.3. Embedding of a ball

For $\sigma \in \mathfrak{S}_n$ and $k \in \{0, \dots, \binom{n}{2}\}$ we denote by $B(\sigma, k)$ the (closed) ball for the Kendall’s tau distance of center σ and radius k , i.e. $B(\sigma, k) = \{\sigma' \in \mathfrak{S}_n \mid d(\sigma, \sigma') \leq k\}$. The following is a direct consequence of Proposition 1.

Lemma 3. For $\sigma \in \mathfrak{S}_n$ and $k \in \{0, \dots, \binom{n}{2}\}$,

$$\phi(\mathfrak{S}_n \setminus B(\sigma, k)) \subset \mathbb{S} \setminus B(\phi(\sigma), 2\sqrt{k+1})$$

5.4. Proof of Theorem 1

We can now prove Theorem 1 by combining the previous results and observations.

Proof of Theorem 1. Let $\mathcal{D}_N \in \mathfrak{S}_n^N$ be a dataset and $\sigma \in \mathfrak{S}_n$ a permutation. By Lemma 2, one has for any $r > 0$,

$$\begin{aligned} \cos(\theta_N(\sigma)) &> \sqrt{1 - \frac{r^2}{4R^2}} \\ &\Rightarrow \min_{x \in \mathbb{S} \setminus B(\phi(\sigma), r)} \theta_N(x) > \theta_N(\sigma). \end{aligned}$$

We take $r = 2\sqrt{k+1}$. The left-hand term becomes $\cos(\theta_N(\sigma)) > \sqrt{1 - \frac{k+1}{R^2}}$, which is the condition in Theorem 1. The right-hand term becomes:

$$\min_{x \in \mathbb{S} \setminus B(\phi(\sigma), 2\sqrt{k+1})} \theta_N(x) > \theta_N(\sigma),$$

which implies by Lemma 3 that

$$\min_{\sigma' \in \mathfrak{S}_n \setminus B(\sigma, k)} \theta_N(\sigma') > \theta_N(\sigma).$$

This means that for all $\sigma' \in \mathfrak{S}_n$ with $d(\sigma, \sigma') > k$, $\theta_N(\sigma') > \theta_N(\sigma)$. Now, by Lemma 1, any Kemeny consensus σ^* necessarily satisfies $\theta_N(\sigma^*) \leq \theta_N(\sigma)$. One therefore has $d(\sigma, \sigma^*) \leq k$, and the proof is concluded. \square

6. Numerical experiments

In this section we study the tightness of the bound in Theorem 1 and the applicability of Method 1 through numerical experiments.

6.1. Tightness of the bound

Recall that we denote by n the number of alternatives, by $\mathcal{D}_N \in \mathfrak{S}_n^N$ any dataset, by r any voting rule, and by $r(\mathcal{D}_N)$ a consensus of \mathcal{D}_N given by r . For ease of notation convenience, we assume that \mathcal{K}_N contains a single consensus (otherwise we pick one randomly as we do in all experiments). The approximation efficiency of r to Kemeny’s rule is exactly measured by $d(r(\mathcal{D}_N), \mathcal{K}_N)$. Applying our method with $r(\mathcal{D}_N)$ would return an upper bound for $d(r(\mathcal{D}_N), \mathcal{K}_N)$, that is:

$$d(r(\mathcal{D}_N), \mathcal{K}_N) \leq k_{\min}.$$

Notably here we are not interested in studying the approximation efficiency of a particular voting rule, but we are rather interested in studying the approximation efficiency specific to our method indicated by the tightness of the bound, i.e.,

$$s(r, \mathcal{D}_N, n) := k_{\min} - d(r(\mathcal{D}_N), \mathcal{K}_N).$$

In other words, $s(r, \mathcal{D}_N, n)$ quantifies how confident we are when we use k_{\min} to “approximate” the approximation efficiency $d(r(\mathcal{D}_N), \mathcal{K}_N)$ of r to Kemeny’s rule on a given dataset \mathcal{D}_N . The smaller $s(r, \mathcal{D}_N, n)$ is, the better our method works when it is combined with the voting rule r to pinpoint the Kemeny consensus on a given dataset \mathcal{D}_N . Note that our notation stresses on the fact that s depends typically on (r, \mathcal{D}_N, n) .

We empirically investigate the efficiency of our proposed method by experimenting $s(r, \mathcal{D}_N, n)$ with various voting rules r , on different datasets \mathcal{D}_N , implicitly involving n as well. For that purpose, in each experiment we test six prevalent voting rules plus one negative-control method as approximate candidates to Kemeny’s rule: three scoring rules that are Borda Count, k -approval, Copeland; two algorithmic approaches that are QuickSort and Pick-a-Perm; one statistical approach based on Plackett-Luce ranking model; one baseline method serving a negative control that is Pick-a-Random where a random permutation is picked from \mathfrak{S}_n according to uniform law (independent from the dataset \mathcal{D}_N). Details of the voting rules are found in the supplementary material.

We first look at the the effect of different voting rules r on $s(r, \mathcal{D}_N, n)$ with the APA dataset. In the 1980 American Psychological Association (APA) presidential election, voters were asked to rank $n = 5$ candidates in order of preference and a total of $N = 5738$ complete ballots were

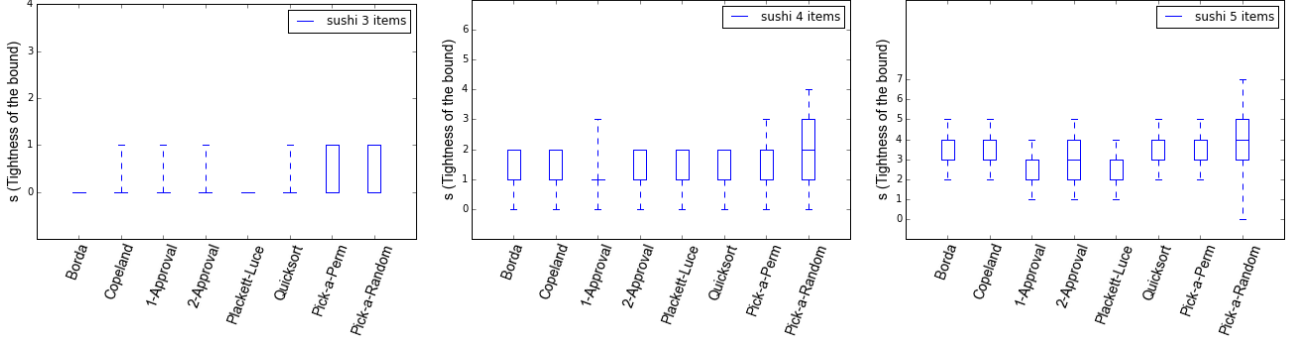


Figure 4. Boxplot of $s(r, \mathcal{D}_N, n)$ over sampling collections of datasets shows the effect from different size of alternative set n with restricted sushi datasets ($n = 3; 4; 5, N = 5000$).

reported. With the original collection of ballots introduced by Diaconis (1989), We created 500 bootstrapped pseudo-samples following Popova (2012). As shown in Figure 5, $s(r; \mathcal{D}_N, n)$ varies across different voting rules and our method works typically well combined with Borda Count or Plackett-Luce, a phenomenon that constantly occurs in many experiments. For example for Borda Count the median tightness being 3 means that our method provides a bound that tolerates an approximation within a Kendall’s tau distance up to 3. We also observe that on the contrary, the boxplot of Pick-a-Random always shows a wider range and larger median as expected.

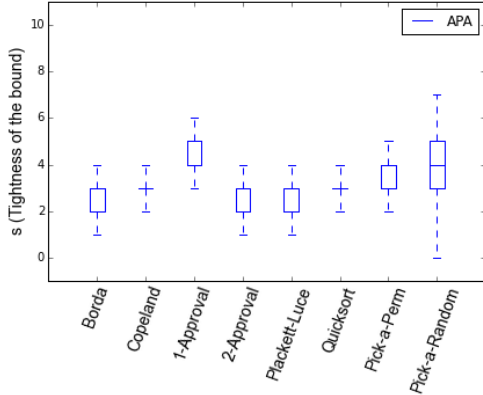


Figure 5. Boxplot of $s(r, \mathcal{D}_N, n)$ over sampling collections of datasets shows the effect from different voting rules r with 500 bootstrapped pseudo-samples of the APA dataset ($n = 5, N = 5738$).

The effect of datasets \mathcal{D}_N on the measure $s(\mathcal{D}_N; r, n)$ is tested with the Netflix data provided by Mattei et al. (2012). We set $n = 3$ the number of ranked alternatives and take two types of data with distinct characteristics to contrast their impact: we took the 100 datasets with a Condorcet winner and randomly selected 100 datasets from those with no Condorcet winner. The rationale for this experiment is that Kemeny’s rule is a Condorcet method, i.e., Kemeny rule always yields a Condorcet winner if it exists. Therefore we suppose that the efficiency of our method should

also depend on this particular social characteristic present in data. As expected, it is interesting to note the clear difference shown by the two types of data shown by Figure 6. In words, our method is more efficient in case that a Condorcet winner is present in the dataset than the other case that a Condorcet winner is absent in the sense that s is generally smaller in the former case.

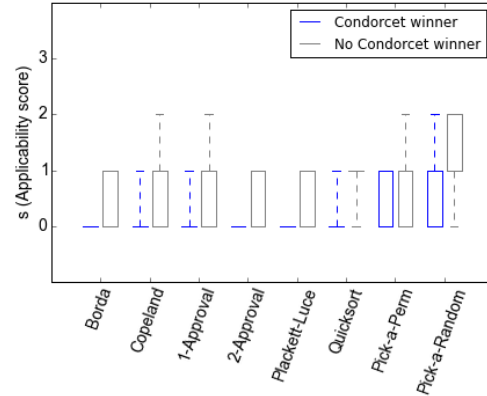


Figure 6. Boxplot of $s(r, \mathcal{D}_N, n)$ over sampling collections of datasets shows the effect from datasets \mathcal{D}_N . 100 Netflix datasets with the presence of Condorcet winner and 100 datasets with no Condorcet winner ($n = 4$ and N varies for each sample).

We finally study how the $s(n; r, \mathcal{D}_N)$ grows with the size of the alternative set n using the sushi dataset found in Kamishima (2003), originally provided as a dataset of $N = 5000$ full rankings of 10 sushi dishes. As evaluating s requires exact Kemeny consensus which can quickly become intractable when n is large, we strict in this study the number of sushi dishes n to be relatively small, and generate collections of datasets, indexed by combinations of n sushi dishes out of $\{1, \dots, 10\}$, by counting the total occurrences of such order present in the original dataset. For example, when $n = 3$ we have a total of $\binom{10}{3} = 120$ different combinations of alternatives (hence 120 collections of datasets) each generated by counting the total occurrences of preference orders of individuals restricted to these 3 alternatives. Therefore we have a total of 120; 210; 252 datasets respec-

tively for $n = 3; 4; 5$. Figure 4 shows that $s(r, \mathcal{D}_N, n)$ increases as n grows, a trend that is dominant and consistent across all voting rules. Since the maximal distance $\binom{n}{2}$ in \mathfrak{S}_n grows quadratically with respect to n , an interesting question would remain to specify explicitly the dependency of k_{min} on n , or the dependency of $s(r, \mathcal{D}_N, n)$ on n , for a given voting rule.

6.2. Applicability of the method

We have so far focused on small n ($n \leq 5$) case, and verified that our method is efficient in using k_{min} to approximate $d(r(\mathcal{D}_N), \mathcal{K}_N)$. We are now mostly interested in the usefulness of our method when k_{min} is directly combined with voting rules in pinpointing Kemeny consensus \mathcal{K}_N particularly when n is large. Now we employ our method by using k_{min} for each dataset to upper bound the approximation performance of $r(\mathcal{D}_N)$ to Kemeny’s rule. Moreover, suppose that we are still interested in finding the exact Kemeny consensus despite a good approximation $r(\mathcal{D}_N)$. Once we have computed an approximated ranking $r(\mathcal{D}_N)$ and k_{min} is identified via our method, the search scope for the exact Kemeny consensus can be narrowed down to those permutations within a distance of k_{min} to $r(\mathcal{D}_N)$. Notably Wang et al. (2013, Lemma 1) proved that the total number of such permutations in \mathfrak{S}_n is upper bounded by $\binom{n+k_{min}-1}{k_{min}}$ which is usually much smaller than $|\mathfrak{S}_n| = n!$.

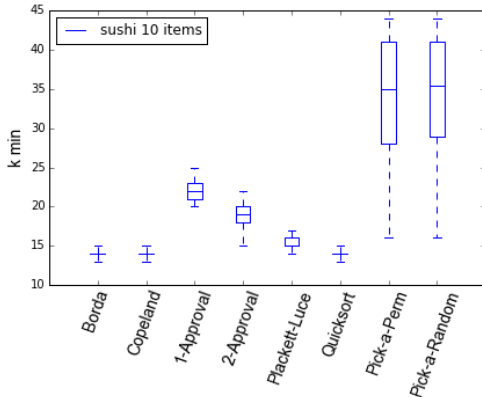


Figure 7. Boxplot of k_{min} over 500 bootstrapped pseudo-samples of the sushi dataset ($n = 10, N = 5000$).

We took the original sushi dataset consisting of $N = 5000$ individual votes on $n = 10$ sushi dishes and created 500 bootstrapped pseudo-samples following the same empirical distribution. Note that k_{min} should also depend on (r, \mathcal{D}_N, n) . Since our bound is established in general with any $\sigma \in \mathfrak{S}_n$ and does take into consideration the approximation efficiency of specific voting rules to Kemeny’s rule, the predicted k_{min} should significantly rely on the approximate voting rules utilized and should be biased more in favor to voting rules with good approximation to Kemeny’s rule since k_{min} can never be inferior to $d(r(\mathcal{D}_N), \mathcal{K}_N)$. As

shown in Figure 7, Pick-a-Random and Pick-a-Perm typically performs poorly, but this is largely due to the fact that the two voting rules are too naive to well approximate Kemeny’s rule *per se*. On the contrary, we observe that Borda, Copeland and Quicksort combined with our method best pinpoint Kemeny consensus with k_{min} of a median distance 14. This further means that in order to obtain all the exact Kemeny consensus now, on average we need to search through at most $\binom{10+14-1}{14} = 817,190$ permutations instead of $10! = 3,628,800$ permutations, where 77% of permutations in \mathfrak{S}_{10} are removed from consideration.

7. Conclusion and discussion

In this paper, we have established a theoretical result that allows to control the Kendall’s tau distance between a permutation and the Kemeny consensus of any dataset. In practice, this provides a simple and general method to predict, for any ranking aggregation procedure, how close its output on a dataset is from the Kemeny consensus. From a broader perspective, it constitutes a novel approach to apprehend the complexity of Kemeny aggregation.

Our results rely on the geometric properties of the Kemeny embedding. Though it has rarely been used in the literature, it provides a powerful framework to analyze Kemeny aggregation. We therefore believe that it could lead to other profound results. In particular we deem that an analysis of how the embeddings of the permutation spread on the sphere could lead to a finer condition in Theorem 1 which is left as future work.

Another interesting direction would certainly be to extend our method to rank aggregation from partial orders, such as pairwise comparisons or top- k rankings. Two main approaches can be followed. In the first one, a partial order would be identified with the set $S \subset \mathfrak{S}_n$ of its linear extensions and its distance to a permutation $\sigma \in \mathfrak{S}_n$ defined by the average $(1/|S|) \sum_{\sigma' \in S} d(\sigma, \sigma')$. The Kemeny embedding would then naturally be extended to S as $(1/|S|) \sum_{\sigma' \in S} \phi(\sigma')$, the barycenter of embeddings of its linear extensions. In the second approach, one would see a partial order as a collection of pairwise comparisons $\{i_1 \succ j_1, \dots, i_m \succ j_m\}$ and define its distance to a permutation $\sigma \in \mathfrak{S}_n$ by the average number of pairwise disagreements $(1/m) \sum_{r=1}^m \mathbb{I}\{\sigma(i_r) > \sigma(j_r)\}$. The Kemeny embedding would then naturally be extended to $\{i_1 \succ j_1, \dots, i_m \succ j_m\}$ as the embedding of any linear extension σ where the coordinate on $\{i, j\}$ is put equal to 0 if $\{i, j\}$ does not appear in the collection. In both cases, our approach would apply with slight changes to exploit the related geometrical properties.

Acknowledgments

Yunlong Jiao was funded by the European Union 7th Framework Program through the Marie Curie ITN MLPM grant No 316861, and by the European Research Council grant ERC-SMAC-280032.

References

- Ailon, N., Charikar, M., and Newman, A. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM (JACM)*, 55(5):23:1–23:27, 2008.
- Ali, A. and Meilă, M. Experiments with kemeny ranking: What works when? *Mathematical Social Sciences*, 64(1):28–40, 2012.
- Arrow, K. J. A difficulty in the concept of social welfare. *The Journal of Political Economy*, pp. 328–346, 1950.
- Barthelemy, J. P. and Monjardet, B. The median procedure in cluster analysis and social choice theory. *Mathematical Social Sciences*, 1:235–267, 1981.
- Bartholdi, J. J., Tovey, C. A., and Trick, M. A. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989.
- Betzler, N., Fellows, M. R., Guo, J., Niedermeier, R., and Rosamond, F. A. Fixed-parameter algorithms for kemeny scores. In *Algorithmic Aspects in Information and Management*, pp. 60–71. Springer, 2008.
- Betzler, N., Fellows, M. R., Guo, J., Niedermeier, R., and Rosamond, F. A. How similarity helps to efficiently compute kemeny rankings. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 657–664. International Foundation for Autonomous Agents and Multiagent Systems, 2009.
- Blin, G., Crochemore, M., Hamel, S., and Vialette, S. Median of an odd number of permutations. *Pure Mathematics and Applications*, 21(2):161–175, 2011.
- Borda, J. C. Mémoire sur les élections au scrutin. 1781.
- Brandt, F., Brill, M., Hemaspaandra, E., and Hemaspaandra, L. A. Bypassing combinatorial protections: Polynomial-time algorithms for single-peaked electorates. *Journal of Artificial Intelligence Research*, pp. 439–496, 2015.
- Cohen, W. W., Schapire, R. E., and Singer, Y. Learning to order things. *Journal of Artificial Intelligence Research*, 10(1):243–270, may 1999.
- Condorcet, N. *Essai sur l’application de l’analyse à la probabilité des décisions rendues à la pluralité des voix*. L’imprimerie royale, 1785.
- Conitzer, V., Davenport, A., and Kalagnanam, J. Improved bounds for computing kemeny rankings. In *AAAI*, volume 6, pp. 620–626, 2006.
- Coppersmith, D., Fleischer, L., and Rudra, A. Ordering by weighted number of wins gives a good ranking for weighted tournaments. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithm, SODA ’06*, pp. 776–782, 2006.
- Cornaz, D., Galand, L., and Spanjaard, O. Kemeny elections with bounded single-peaked or single-crossing width. In *IJCAI*, volume 13, pp. 76–82. Citeseer, 2013.
- Davenport, A. and Kalagnanam, J. A computational study of the kemeny rule for preference aggregation. In *AAAI*, volume 4, pp. 697–702, 2004.
- Davenport, A. and Lovell, D. Ranking pilots in aerobatic flight competitions. Technical report, IBM Research Report RC23631 (W0506-079), TJ Watson Research Center, NY, 2005.
- Deng, K., Han, S., Li, K. J., and Liu, J. S. Bayesian aggregation of order-based rank data. *Journal of the American Statistical Association*, 109(507):1023–1039, 2014.
- Desarkar, M. S., Sarkar, S., and Mitra, P. Preference relations based unsupervised rank aggregation for metasearch. *Expert Systems with Applications*, 49:86 – 98, 2016.
- Diaconis, P. A generalization of spectral analysis with application to ranked data. *The Annals of Statistics*, pp. 949–979, 1989.
- Diaconis, P. and Graham, R. L. Spearman’s footrule as a measure of disarray. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 262–268, 1977.
- Dwork, C., Kumar, R., Naor, M., and Sivakumar, D. Rank aggregation methods for the Web. In *Proceedings of the 10th International WWW conference*, pp. 613–622, 2001.
- Freund, D. and Williamson, D. P. Rank aggregation: New bounds for mxc. *CoRR*, abs/1510.00738, 2015.
- Jiang, X., Lim, L.-H., Yao, Y., and Ye, Y. Statistical ranking and combinatorial Hodge theory. *Mathematical Programming*, 127(1):203–244, 2011.
- Jiao, Y. and Vert, J.-P. The kendall and mallows kernels for permutations. In Blei, D. and Bach, F. (eds.), *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, pp. 1935–1944, 2015.

- Kamishima, T. Nantonac collaborative filtering: recommendation based on order responses. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 583–588. ACM, 2003.
- Kemeny, J. G. Mathematics without numbers. *Daedalus*, 88:571–591, 1959.
- Kolde, R., Laur, S., Adler, P., and Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics*, 28(4):573–580, 2012.
- Lahaie, S. and Shah, N. Neutrality and geometry of mean voting. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 333–350. ACM, 2014.
- Li, H. Learning to rank for information retrieval and natural language processing. *Synthesis Lectures on Human Language Technologies*, 7(3):1–121, 2014.
- Mattei, N., Forshee, J., and Goldsmith, J. An empirical study of voting rules and manipulation with large datasets. In *Proceedings of COMSOC*. Citeseer, 2012.
- Patel, T., Telesca, D., Rallo, R., George, S., Xia, T., and Nel, A. E. Hierarchical rank aggregation with applications to nanotoxicology. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(2):159–177, 2013.
- Popova, A. The robust beauty of apa presidential elections: an empty-handed hunt for the social choice conundrum. Master’s thesis, University of Illinois at Urbana-Champaign, 2012. URL <http://hdl.handle.net/2142/34562>.
- Procaccia, A. D., Reddi, S. J., and Shah, N. A maximum likelihood approach for selecting sets of alternatives. *CoRR*, 2012.
- Renda, M. E. and Straccia, U. Web metasearch: rank vs. score based rank aggregation methods. In *Proceedings of the 2003 ACM symposium on Applied computing*, pp. 841–846. ACM, 2003.
- Saari, D. G. and Merlin, V. R. A geometric examination of kemeny’s rule. *Social Choice and Welfare*, 17(3):403–438, 2000.
- Schalekamp, F. and van Zuylen, A. Rank aggregation: Together we’re strong. In *Proceedings of the Meeting on Algorithm Engineering & Experiments*, pp. 38–51. Society for Industrial and Applied Mathematics, 2009.
- Sibony, E. Borda count approximation of kemeny’s rule and pairwise voting inconsistencies. In *NIPS-2014 Workshop on Analysis of Rank Data: Confluence of Social Choice, Operations Research, and Machine Learning*. Curran Associates, Inc., 2014.
- Stanley, R. P. *Enumerative Combinatorics*. Wadsworth Publishing Company, Belmont, CA, USA, 1986. ISBN 0-534-06546-5.
- Van Zuylen, A. and Williamson, D. P. Deterministic algorithms for rank aggregation and other ranking and clustering problems. In *Approximation and Online Algorithms*, pp. 260–273. Springer, 2007.
- Wang, D., Mazumdar, A., and Wornell, G. W. A rate-distortion theory for permutation spaces. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pp. 2562–2566. IEEE, 2013.
- Xia, L. Generalized decision scoring rules: Statistical, computational, and axiomatic properties. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC ’15, pp. 661–678, New York, NY, USA, 2015. ACM.
- Young, H. P. and Levenglick, A. A consistent extension of condorcet’s election principle. *SIAM Journal on applied Mathematics*, 35(2):285–300, 1978.
- Zamani, H., Shakery, A., and Moradi, P. Regression and learning to rank aggregation for user engagement evaluation. In *Proceedings of the 2014 Recommender Systems Challenge*, pp. 29. ACM, 2014.
- Zwicker, W. S. Consistency without neutrality in voting rules: When is a vote an average? *Mathematical and Computer Modelling*, 48(9):1357–1373, 2008.