Borda count approximation of Kemeny's rule and pairwise voting inconsistencies

Eric Sibony
LTCI UMR No. 5141, Telecom ParisTech/CNRS
Institut Mines-Telecom
Paris, 75013, France
eric.sibony@telecom-paristech.fr

Abstract

We establish an inequality that relates the cost of the Borda Count approximation of Kemeny's Rule to a quantitative measure of the amount of pairwise inconsistencies in the data. This result provides a deeper understanding of the connection between the Borda Count and Kemeny's Rule as well as new theoretical guarantees for the application of this approximation procedure in practice. Its establishment relies on a novel framework for consensus ranking approximation that opens the way to further similar results.

1 Introduction

Thoroughly studied in the context of social choice, ranking aggregation has also attracted much attention in the machine learning literature in the last years, due to its numerous modern applications (see [8], [10], [12], [11], [13] , [2] or [15] for instance). If $[\![n]\!] = \{1, \dots, n\}$ designates a set of items or candidates and $\sigma_1, \dots, \sigma_N$ are ordered lists (seen as permutations) on $[\![n]\!]$ representing votes of electors or results of different search engines, ranking aggregation consists in finding a permutation σ^* that best "summarizes" the collection $\sigma_1, \dots, \sigma_N$. One widely used rationale is to formulate this problem as the search of a "consensus ranking" σ^* that minimizes the sum of the distances to the σ_t's, *i.e.* a solution to the minimization problem $\min_{\sigma \in \mathfrak{S}_n} \sum_{t=1}^{N} d(\sigma, \sigma_t)$, where d is a given metric on $\mathfrak{S}_n$, the set of permutations on $[\![n]\!]$ (such an element always exists, as $\mathfrak{S}_n$ is finite, but is not necessarily unique).

The most studied consensus is certainly the Kemeny's Rule, when the metric d is equal to the Kendall's tau distance (see section 3 for a definition). Despite its many good properties, the computation of such a consensus is however NP-hard, see [3] and [8], and many aggregation procedures, or voting rules, have thus been proposed in the literature as approximations schemes. The Borda Count (see section 4 for a definition) has especially been studied. In particular it has been shown that it is is a 5-approximation of Kemeny's rule (see [6]), that it recovers some part of the Kemeny consensus under noise assumptions (see [19] and [14]), and that it provides a good trade-off between accuracy and computational complexity for several datasets (see [1]).

Results from the *geometric voting theory*, introduced and developed by Saari (see [17] for instance), have also provided a deeper understanding of the differences between the outputs of the Borda Count and Kemeny's Rule, showing in particular that they are due to the local inconsistencies between pairwise votes in the data, see [18], [4] or [7]. Similarly, in the *HodgeRank* framework introduced in [9], it was shown that the Borda Count corresponds to an $l2$ projection of the data on the space of "gradient flows" (localizing the global consistencies) orthogonal to the space of "divergence-free flows" (localizing local pairwise inconsistencies).

The main contribution of this paper is to establish a new theoretical guarantee for the approximation of the Kemeny's Rule by the Borda Count (theorem 4), given by an inequality of the form:

$\sum_{t=1}^{N} d(\sigma^{BC}, \sigma_t) - \sum_{t=1}^{N} d(\sigma^*, \sigma_t) \leq C(\sigma_1, \ldots, \sigma_N)$, where σ^{BC} is output by the Borda Count for $\sigma_1, \ldots, \sigma_N$ and $C(\sigma_1, \ldots, \sigma_N)$ is a quantitative measure of the amount of local pairwise inconsistencies in the data. Beyond deepening the explanation of the differences between the Borda Count and Kemeny's Rule, this result provides new theoretical guarantees for practical empirical work.

2 Notations and General results

Instead of dealing directly with the collection of permutations $(\sigma_1, \ldots, \sigma_N)$ we adopt the point of view of the geometric voting theory and consider the function $p : \mathfrak{S}_n \to \mathbb{N}$ that counts the number of occurrences of each $\pi \in \mathfrak{S}_n$ in the collection, *i.e.* $p(\pi) = |\{1 \leq t \leq N \mid \sigma_t = \pi\}|$ for $\pi \in \mathfrak{S}_n$. A consensus ranking of the collection $\sigma_1, \ldots, \sigma_N \in \mathfrak{S}_n$ with respect to a metric d is then a permutation σ^* such that

$$\sum_{\pi \in \mathfrak{S}_n} d(\sigma^*, \pi) p(\pi) = \min_{\sigma \in \mathfrak{S}_n} \sum_{\pi \in \mathfrak{S}_n} d(\sigma, \pi) p(\pi). \tag{1}$$

The function p is called a *voting profile*, and this reformulation shows that the consensus ranking problem can be directly defined for a voting profile or even more generally for any real-valued function p on $\mathfrak{S}_n$ (that we still call a profile). Let $L(\mathfrak{S}_n) = \{f : \mathfrak{S}_n \to \mathbb{R}\}$ be the space of real valued-functions on $\mathfrak{S}_n$, equipped with its canonical inner product $\langle f, g \rangle = \sum_{\sigma \in \mathfrak{S}_n} f(\sigma) g(\sigma)$. For $p \in L(\mathfrak{S}_n)$ and d a metric on $\mathfrak{S}_n$, we define the cost function $\mathcal{R}_{d,p}(\sigma) = \sum_{\pi \in \mathfrak{S}_n} d(\sigma, \pi) p(\pi)$ for $\sigma \in \mathfrak{S}_n$, and denote its minimum by $\mathcal{R}^*_{d,p} = \min_{\sigma \in \mathfrak{S}_n} \mathcal{R}_{d,p}(\sigma)$. The consensus rankings of p with respect to d are then the elements of $\mathcal{C}_d(p) = \{\sigma^* \in \mathfrak{S}_n \mid \mathcal{R}(\sigma^*) = \mathcal{R}^*_{d,p}\}$. Notice that for $\lambda > 0$ and $\mu \in \mathbb{R}$, $\mathcal{C}_d(\lambda p + \mu) = \mathcal{C}_d(p)$.

The major advantage of the point of view of the geometric voting theory is to allow the use of tools from linear algebra. For a metric d on $\mathfrak{S}_n$, we denote by $d_{max} = \max_{\sigma, \pi \in \mathfrak{S}_n} d(\sigma, \pi)$ its diameter and we consider the $n! \times n!$ matrix T_d defined by $T_d(\sigma, \pi) = d_{max} - d(\sigma, \pi)$ for $\sigma, \pi \in \mathfrak{S}_n$. The cost function is then equal to $\mathcal{R}_{d,p}(\sigma) = C - T_d p(\sigma)$ with $C \in \mathbb{R}$ a constant, and the set of consensus rankings of the profile p according to the metric d is simply equal to the modes of $T_d p$:

$$\mathcal{C}_d(p) = \{\sigma^* \in \mathfrak{S}_n \mid T_d p(\sigma^*) = \max_{\sigma \in \mathfrak{S}_n} T_d p(\sigma)\}. \tag{2}$$

The following theorem establishes a general inequality that we will use to prove our main result.

Theorem 1. *Let d be a metric on $\mathfrak{S}_n$ and $p, q \in L(\mathfrak{S}_n)$. Then for any $\sigma \in \mathcal{C}_d(q)$,*

$$\mathcal{R}_{d,p}(\sigma) - \mathcal{R}^*_{d,p} \ \leq \ 2 \| T_d (p - q) \|_\infty.$$

Proof. For $\sigma \in \mathcal{C}_d(q)$, $T_d q(\sigma) = \max_{\pi \in \mathfrak{S}_n} T_d q(\pi)$ so $\mathcal{R}_{d,p}(\sigma) - \mathcal{R}^*_{d,p} = \max_{\pi \in \mathfrak{S}_n} T_d p(\pi) - \max_{\pi \in \mathfrak{S}_n} T_d q(\pi) + T_d(q - p)(\sigma) \leq \max_{\pi \in \mathfrak{S}_n} T_d(p - q)(\pi) + \max_{\pi \in \mathfrak{S}_n} T_d(q - p)$. □

The reformulation (2) of the consensus ranking problem as well as theorem 1 highlight the fact that for a given metric d, the consensus ranking problem related to a profile p is fully characterized by $T_d p$ and not p. This motivates the following definition, already considered in [17] and [7].

Definition 1 (Effective space)**.** The effective space of a linear transform T on a Euclidean space V is $(\ker T)^\perp$, the orthogonal supplementary of its null space.

In the present context, T_d is symmetric so $(\ker T_d)^\perp = \operatorname{Im} T_d$, and T_d is diagonalizable. Concretely, if we denote by $E_1, \ldots, E_r$ its eigenspaces respectively associated with its distinct non-null eigenvalues $\lambda_1, \ldots, \lambda_r$, then $L(\mathfrak{S}_n) = \ker T_d \oplus \bigoplus_{i=1}^{r} E_i$, where the sum is orthogonal, and if $p = p_{\ker T_d} + \sum_{i=1}^{r} p_i$ is the corresponding linear decomposition of $p \in L(\mathfrak{S}_n)$, $T_d p = \sum_{i=1}^{r} \lambda_i p_i$. Many results on the consensus ranking problem thus rely on the eigenstructure of the matrix T_d. The following lemma is one illustration (its proof is straightforward and left to the reader).

Lemma 1. *If $T_d p = \lambda p$ with $\lambda > 0$, then $\mathcal{C}_d(p) = \{\sigma^* \in \mathfrak{S}_n \mid p(\sigma^*) = \max_{\sigma \in \mathfrak{S}_n} p(\sigma)\}$.*

3 Spectral decomposition of the Kemeny consensus effective space

Here and in the sequel, d is the Kendall's tau distance on $\mathfrak{S}_n$, defined for $\pi, \sigma \in \mathfrak{S}_n$ as the number of discordant pairs between π and σ: $d(\pi, \sigma) = |\{1 \leq i < j \leq n \mid (\sigma(j) - \sigma(i))(\pi(j) - \pi(i)) < 0\}|$. One clearly has $d_{max} = \binom{n}{2}$. It is well known that the Kendall's tau distance is the metric of the Cayley graph on $\mathfrak{S}_n$ generated by adjacent transpositions, that is to say the graph with $\mathfrak{S}_n$ as vertices where there is an edge between $\sigma \in \mathfrak{S}_n$ and $\pi \in \mathfrak{S}_n$ if $\sigma\pi^{-1} = (i\ i{+}1)$ with $1 \leq i < n$. In [16], the eigenstructure of the distance matrix D of this graph, which is exactly defined by $D(\sigma, \pi) = d(\sigma, \pi)$, was fully determined. Since $T_d = \binom{n}{2} J - D$ where J is the $n! \times n!$ with only ones, the following theorem is a direct consequence of theorem 39 in [16] (for the A_{n-1} type, see the Table 5).

Theorem 2 ([16])**.** *The eigenvalues of T_d are $n!/2\binom{n}{2}$, $(n+1)!/6$, $n!/6$ and 0, with respective multiplicities 1, $n-1$, $\binom{n-1}{2}$ and $n! - \binom{n}{2} - 1$.*

Theorem 2 shows in particular that the effective space of T_d has dimension $\binom{n}{2} + 1$. We introduce some more notations to fully characterize it. We denote by $\mathbb{1}_S$ the indicator function of a subset $S \subset \mathfrak{S}_n$. For $i, j \in [\![n]\!]$, $i \neq j$, we define the function $K_{i,j}(\sigma) = \operatorname{sign}(\sigma(j) - \sigma(i))$ for $\sigma \in \mathfrak{S}_n$, where $\operatorname{sign}(x) = x/|x|$ for $x \in \mathbb{R} \setminus \{0\}$, *i.e.* $K_{i,j} = \mathbb{1}_{\{\sigma(i)<\sigma(j)\}} - \mathbb{1}_{\{\sigma(j)<\sigma(i)\}}$, where $\mathbb{1}_{\{\sigma(i)<\sigma(j)\}}$ is a shortcut for $\mathbb{1}_{\{\sigma \in \mathfrak{S}_n \mid \sigma(i)<\sigma(j)\}}$. Since for $1 \leq i < j \leq n$, $K_{j,i} = -K_{i,j}$, $\operatorname{span}(K_{i,j})_{1 \leq i,j \leq n, i \neq j} = \operatorname{span}(K_{i,j})_{1 \leq i < j \leq n}$, and we denote this space by $\mathcal{K}_n$.

Proposition 1. *For $p \in L(\mathfrak{S}_n)$,*

$$T_d p = \sum_{1 \leq i \neq j \leq n} \left\langle p, \mathbb{1}_{\{\sigma(i)<\sigma(j)\}} \right\rangle \mathbb{1}_{\{\sigma(i)<\sigma(j)\}} = \frac{1}{2}\binom{n}{2} \langle p, \mathbb{1}_{\mathfrak{S}_n} \rangle \mathbb{1}_{\mathfrak{S}_n} + \frac{1}{2} \sum_{1 \leq i < j \leq n} \langle p, K_{i,j} \rangle K_{i,j}.$$

In particular, $\dim \mathcal{K}_n = \binom{n}{2}$ *and* $(\ker T_d)^\perp = \operatorname{Im} T_d = \mathbb{R}\mathbb{1}_{\mathfrak{S}_n} \oplus \mathcal{K}_n$.

Proof. The first equality is obtained by noticing that for $\pi \in \mathfrak{S}_n$,

$$\sum_{1 \leq i \neq j \leq n} \left\langle p, \mathbb{1}_{\{\sigma(i)<\sigma(j)\}} \right\rangle \mathbb{1}_{\{\sigma(i)<\sigma(j)\}}(\pi) = \sum_{\sigma \in \mathfrak{S}_n} p(\sigma) \sum_{1 \leq i \neq j \leq n} \mathbb{I}\{\sigma(i) < \sigma(j)\}\mathbb{I}\{\pi(i) < \pi(j)\},$$

where $\mathbb{I}\{\mathcal{E}\}$ denotes the indicator function of the event $\mathcal{E}$, and the second equality only uses that $\mathbb{1}_{\{\sigma(i)<\sigma(j)\}} + \mathbb{1}_{\{\sigma(j)<\sigma(i)\}} = \mathbb{1}_{\mathfrak{S}_n}$ for $i \neq j$. It implies that $(\ker T_d)^\perp = \operatorname{Im} T_d \subset \mathbb{R}\mathbb{1}_{\mathfrak{S}_n} \oplus \mathcal{K}_n$ and as $\dim \mathcal{K}_n \leq \binom{n}{2}$ and $\dim(\ker T_d)^\perp = \binom{n}{2} + 1$, the proof is concluded. □

Remark 1. Proposition 1 shows in particular that the computation of $T_d p$ only involves the quantities $\left\langle p, \mathbb{1}_{\{\sigma(i)<\sigma(j)\}} \right\rangle$. Interpreted in the multiresolution analysis framework introduced in [5], this means that all the information of p of scale higher than 2 is irrelevant to the computation of the Kemeny consensus of p.

An explicit construction of the eigenspaces of T_d is also provided in [16] but the notations are very far from ours. Furthermore, it happens that these eigenspaces are equal to well studied spaces in the geometric theory of voting.

Definition 2. For $i, j \in [\![n]\!]$ with $i \neq j$, let B_i and $C_{i,j}$ be the functions defined for $\sigma \in \mathfrak{S}_n$ by

$$B_i(\sigma) = n + 1 - 2\sigma(i) \qquad \text{and} \qquad C_{i,j}(\sigma) = n \operatorname{sign}(\sigma(j) - \sigma(i)) - 2(\sigma(j) - \sigma(i)),$$

and define the spaces $\mathcal{B}_n = \operatorname{span}(B_i)_{1 \leq i \leq n}$ and $\mathcal{C}_n = \operatorname{span}(C_{i,j})_{1 \leq i,j \leq n, i \neq j}$.

The space $\mathcal{B}_n$ is called the Borda component in [17], it is shown to have dimension $n-1$ and to admit $(B_i)_{1 \leq i \leq n-1}$ as a basis. The space $\mathcal{C}_n$ is called the Condorcet component in [7], it is shown to have dimension $\binom{n-1}{2}$ and to admit $(C_{i,j})_{1 \leq i < j \leq n-1}$ as a basis. The following lemma shows that both $\mathcal{B}_n$ and $\mathcal{C}_n$ are subspaces of $\mathcal{K}_n$. Its proof is only technical and left to the reader.

Lemma 2. *For $i, j \in [\![n]\!]$, $i \neq j$, $B_i = \sum_{k \neq i} K_{i,k}$ and $C_{i,j} = \sum_{k \notin \{i,j\}} (K_{i,j} + K_{j,k} + K_{k,i})$.*

The following theorem makes explicit the eigenspaces of T_d, it is the major result of this section.

Theorem 3. *The eigenspaces of T_d related to the eigenvalues $n!/2\binom{n}{2}$, $(n+1)!/6$ and $n!/6$ are repsectively $\mathbb{R}\mathbb{1}_{\mathfrak{S}_n}$, $\mathcal{B}_n$ and $\mathcal{C}_n$. In particular, the effective space of T_d admits the orthogonal linear decomposition $(\ker T_d)^\perp = \mathbb{1}_{\mathfrak{S}_n} \oplus \mathcal{B}_n \oplus \mathcal{C}_n$.*

Proof. By theorem 2, the proof only requires to show that the spaces $\mathbb{R}\mathbb{1}_{\mathfrak{S}_n}$, $\mathcal{B}_n$ and $\mathcal{C}_n$ are each a subspace of the corresponding eigenspace of T_d. This is done using proposition 1, lemma 2 and the calculation of all the possible values of $\langle K_{i,j}, K_{k,l}\rangle$ for $i, j, k, l \in [\![n]\!]$ with $i \neq j$ and $k \neq l$. □

Proposition 2. *The orthogonal projections of an element $p \in L(\mathfrak{S}_n)$ on $\mathcal{B}_n$ and $\mathcal{C}_n$ are given by*

$$p_{\mathcal{B}_n} = \frac{3}{n(n+1)!}\sum_{i=1}^{n} \langle p, B_i\rangle B_i \qquad \text{and} \qquad p_{\mathcal{C}_n} = \frac{3}{n.n!}\sum_{1\leq i<j\leq n-1} \langle p, K_{i,j} + K_{j,n} + K_{n,i}\rangle C_{i,j}.$$

Proof. We only need to show that the given expressions for $p_{\mathcal{B}_n}$ and $p_{\mathcal{C}_n}$ satisfy respectively $\langle p - p_{\mathcal{B}_n}, B_i\rangle = 0$ for all $i \in [\![n]\!]$ and $\langle p - p_{\mathcal{C}_n}, C_{i,j}\rangle = 0$ for all $i, j \in [\![n]\!]$ with $i \neq j$. This is done through explicit calculations using the formulas of lemma 2. □

4 Borda Count approximation of Kemeny's Rule

For a collection of permutations $\sigma_1, \dots, \sigma_N$, the Borda Count is the voting rule that consists in affecting to the candidate/item $i \in [\![n]\!]$ the score $\sum_{t=1}^{N} \sigma_t(i)$ and then produce a full ranking of the candidates by sorting them in increasing of these scores (notice that this does not define a unique output as some candidates may receive the same score). The Borda Count is generalized to the case of a voting profile $p \in L(\mathfrak{S}_n)$ by replacing the score of the candidate $i \in [\![n]\!]$ by $\sum_{\sigma\in\mathfrak{S}_n} \sigma(i)p(\sigma)$ or equivalently by $-\langle p, B_i\rangle$. The set $BC(p)$ of the outputs of the Borda Count for p is therefore fully characterized by $p_{\mathcal{B}_n}$. The following proposition refines this observation.

Proposition 3. *For $p \in L(\mathfrak{S}_n)$, $BC(p) = \{\sigma^* \in \mathfrak{S}_n \mid p_{\mathcal{B}_n}(\sigma^*) = \max_{\sigma\in\mathfrak{S}_n} p_{\mathcal{B}_n}(\sigma)\} = \mathcal{C}_d(p_{\mathcal{B}_n})$.*

Proof. Using the explicit formula of $p_{\mathcal{B}_n}$ from proposition 2, the first equality comes from the following classic lemma: for $a_1, \dots, a_n \in \mathbb{R}^+$ with $a_n \geq \dots \geq a_1$, the maximum of $\sum_{i=1}^{n} a_i\sigma(i)$ is obtained for $\sigma = id$. The second equality is a direct consequence of lemma 1. □

Proposition 3 says that the outputs of the Borda Count for p are exactly the modes of $p_{\mathcal{B}_n}$ and also exactly the Kemeny consensuses of $p_{\mathcal{B}_n}$. This surely justifies the name of "Borda space" for $\mathcal{B}_n$. The justification of the name "Condorcet space" for $\mathcal{C}_n$ comes from the fact that taking $p_{\mathcal{C}_n}$ into account in Kemeny's rule in addition to $p_{\mathcal{B}_n}$ outputs rankings for which the winners are Condorcet winners (see theorem 3 of [18]). Our result provides a quantitative bound on the difference of costs.

Theorem 4. *For $p \in L(\mathfrak{S}_n)$ and $\sigma^{BC} \in BC(p)$,*

$$\mathcal{R}_{d,p}\left(\sigma^{BC}\right) - \mathcal{R}^*_{d,p} \leq \frac{n!}{3}\|p_{\mathcal{C}_n}\|_\infty \leq \frac{n-2}{n}\sum_{1\leq i<j\leq n-1} |\langle p, K_{i,j} + K_{j,n} + K_{n,i}\rangle|.$$

In particular, $\mathcal{R}_{d,p}\left(\sigma^{BC}\right) = \mathcal{R}^*_{d,p}$ when p is globally consistent in the sense that $p_{\mathcal{C}_n} = 0$.

Proof. By proposition 3, $\sigma^{BC} \in \mathcal{C}_d(p) = \mathcal{C}_d(p + \mu)$ for any $\mu \in \mathbb{R}$. Theorem 1 thus gives $\mathcal{R}_{d,p}\left(\sigma^{BC}\right) - \mathcal{R}^*_{d,p} \leq 2\|T_d(p - p_{\mathcal{B}_n} - \langle p, \mathbb{1}_{\mathfrak{S}_n}\rangle/n!)\|_\infty = 2\|T_d p_{\mathcal{C}_n}\|_\infty$. The proof is concluded using theorem 3, proposition 2 and the observation that $\|C_{i,j}\|_\infty = n-2$ for $i \neq j$. □

As stated in the introduction, the upper bound in theorem 4 is a quantitative measure of the pairwise inconsistencies of p, each term $|\langle p, K_{i,j} + K_{j,n} + K_{n,i}\rangle|$ representing the inconsistency on the triple $\{i, j, n\}$.

5 Conclusion

Using recent results on the spectral structure of the matrix of the Kendall's tau distance, we showed that the excess cost of the Borda Count approximation of Kemeny's Rule is bounded by a quantitative measure of the pairwise inconsistencies present in the data. Not only this result brings a deeper understanding of the relations between the Borda Count and Kemeny's rule but it also provides a new theoretical guarantee for the Borda Count approximation scheme in practice. The novel framework introduced in this paper can furthermore be used to obtain similar results for other consensuses and voting rules.

References

[1] Alnur Ali and Marina Meila. Experiments with kemeny ranking: What works when? *Mathematical Social Sciences*, 64(1):28 – 40, 2012.

[2] Hossein Azari Soufiani, William Chen, David C Parkes, and Lirong Xia. Generalized method-of-moments for rank aggregation. In *Advances in Neural Information Processing Systems 26*, pages 2706–2714, 2013.

[3] J.J. Bartholdi, C.A. Tovey, and M.A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989.

[4] Abhijit Chandra and Sunanda Roy. On removing condorcet effects from pairwise election tallies. *Social Choice and Welfare*, 40(4):1143–1158, 2013.

[5] Stéphan Clémençon, Jérémie Jakubowicz, and Eric Sibony. Multiresolution analysis of incomplete rankings. *ArXiv e-prints*, 2014.

[6] Don Coppersmith, Lisa Fleischer, and Atri Rudra. 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, pages 776–782, 2006.

[7] Karl-Dieter Crisman. The Borda count, the Kemeny rule, and the permutahedron. *Contemporary Mathematics*, 624, 2014.

[8] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the Web. In *Proceedings of the 10th International WWW conference*, pages 613–622, 2001.

[9] Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial Hodge theory. *Math. Program.*, 127(1):203–244, 2011.

[10] Guy Lebanon and John Lafferty. Cranking: Combining rankings using conditional probability models on permutations. In *Proceedings of the 19th International Conference on Machine Learning*, pages 363–370, 2002.

[11] Tyler Lu and Craig Boutilier. The unavailable candidate model: a decision-theoretic view of social choice. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, pages 263–274, 2010.

[12] M. Meila, K. Phadnis, A. Patterson, and J. Bilmes. Consensus ranking under the exponential model. In *Proceedings of UAI'07*, pages 729–734, 2007.

[13] Sahand Negahban, Sewoong Oh, and Devavrat Shah. Iterative ranking from pair-wise comparisons. *CoRR*, 2012.

[14] Ariel D. Procaccia, Sashank Jakkam Reddi, and Nisarg Shah. A maximum likelihood approach for selecting sets of alternatives. *CoRR*, 2012.

[15] Arun Rajkumar and Shivani Agarwal. A statistical convergence perspective of algorithms for rank aggregation from pairwise data. In *Proceedings of the 31st International Conference on Machine Learning*, 2014.

[16] Paul Renteln. The distance spectra of cayley graphs of coxeter groups. *Discrete Mathematics*, 311(89):738 – 755, 2011.

[17] Donald G. Saari. Mathematical structure of voting paradoxes. *Economic Theory*, 15(1):1–53, 2000.

[18] Donald G. Saari and Vincent R. Merlin. A geometric examination of kemeny's rule. *Social Choice and Welfare*, 17(3):403–438, 2000.

[19] H.P. Young. Condorcet's theory of voting. *American Political Science Review*, 82(4):1231–1244, 1988.