Fair Committee Selection under Ordinal Preferences and Limited Cardinal Information

Extended Abstract

Ameet Gadekar
CISPA Helmholtz Center for Information Security
Saarbrücken, Germany

Aristides Gionis
KTH Royal Institute of Technology
Stockholm, Sweden

Suhas Thejaswi
Aalto University
Helsinki, Finland

Sijing Tu
Stanford University
Stanford, USA

ABSTRACT

We study fair k-committee selection under an egalitarian objective. Here, we are given a set of agents partitioned into groups, and each agent reports a complete ranking over all other agents. The task is to select a committee of size k that satisfies minimum representation constraints for each group while minimizing the maximum cost incurred by any agent. Without access to cardinal distances, constant distortion guarantees are not possible for $k \geq 3$ (Burkhardt et al. 2024). In light of this, we model the problem as the ordinal fair k-center problem with limited access to distance queries. In this query-efficient setting, our main result is a 5-distortion algorithms using $O(k \log^2 k)$ queries; along the way, we obtain a 3-distortion algorithm using $O(k^2)$ queries.

KEYWORDS

committee selection; ordinal preferences; metric distortion; ordinal fair k-center; limited queries

ACM Reference Format:
Ameet Gadekar, Aristides Gionis, Suhas Thejaswi, and Sijing Tu. 2026. Fair Committee Selection under Ordinal Preferences and Limited Cardinal Information: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/KFMY7437

1 INTRODUCTION

In many collective decision-making tasks, a group of agents need to select a subset of alternatives to act as representatives. We assume agents and alternatives are embedded in a metric space, where the distance between an agent and an alternative represents the (social) cost the agent experiences when that alternative is selected. Further, each agent provides a complete ranking over the alternatives consistent with these distances—ordinal information, as opposed to explicit distance values *i.e.*, cardinal information. The goal is to choose a committee of size k that minimizes the social cost.

A standard way to assess the performance of algorithms that rely on ordinal preferences is via distortion, *i.e.*, the worst-case ratio between the social cost of the algorithm's outcome and the optimal outcome with full access to cardinal distances [11]. In metric preference settings, constant distortion is achievable in the single-winner selection, even without access to cardinal information [1]. More recently, Pulyassary and Swamy [12] and Burkhardt et al. [3] extended this framework to multi-winner (k-committee) selection for arbitrary k, studying both utilitarian (*e.g.*, k-median, k-means) and egalitarian (k-center) objectives. For k-center, without cardinal information, designing algorithms with constant distortion is not possible for $k \geq 3$. To bypass this impossibility, both works allow limited access to cardinal information via distance queries and design algorithms with constant-factor distortion guarantees.

In our work, we extend this line of research by incorporating both *equity* and *fairness* into committee selection. We capture equity via an egalitarian objective, ensuring no agent is unduly disadvantaged. We model fairness via minimum representation requirements for demographic groups, reflecting the widespread use of quota-based constraints as a mechanism for promoting diverse representation and balancing opportunity [2, 5]. Inline with the work of Burkhardt et al. [3], Cembrano and Shahkarami [4], we consider the setting where agents and alternatives coincide (*i.e.*, agents rank all other agents). When full access to cardinal distances is available, our setting is equivalent to the fair k-center problem [10]. We therefore formulate our setting as the *ordinal fair k-center problem with limited cardinal information* and build upon algorithmic techniques from fair k-center literature [6, 9]. More formally, we study the following problem.

Definition 1.1 (The ordinal fair k-center problem). An instance of the ordinal fair k-center problem is defined on a set U of n points from a metric space (U, d) with unknown d, an integer $k \geq 1$, a collection $\succ_U = \{\succ_u\}_{u \in U}$ of linear orders that is consistent with d, a collection $\mathbb{G} = \{G_1, \ldots, G_t : G_i \subseteq U\}$ of t subsets of data points that form a partition of U, and a vectors of requirements $\vec{\alpha} = \{\alpha_1, \ldots, \alpha_t\}$, where $\alpha_i \geq 0$ corresponds to the requirement of group G_i. A set $S \subseteq U$ of centers is a feasible solution if $|S| = k$ and $\alpha_i \leq |S \cap G_i|$ for all $i \in [t]$. The social cost of a solution S is the maximum distance of any point to S, *i.e.*, $\texttt{cost}(S) = \max_{u \in U} d(u, S)$.

[*] Authors are listed in alphabetical order. This work was carried out while Thejaswi was employed at Max Planck Institute for Software Systems and Tu was at KTH Royal Institute of Technology. Portions of this work was also carried out during visits by Thejaswi and Gadekar to KTH Royal Institute of Technology.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/KFMY7437

The goal of the ordinal fair k-center problem is to find a feasible solution with minimum social cost.

Without loss of generality, we restrict our attention to the setting with exactly k groups, with a requirement of one representative from each group. We show that the general variant reduces to one-per-group setting, and therefore focus on this simpler variant for clarity and brevity (see arXiv version [7] for details of the reduction).

The quality of an algorithm that rely on ordinal preferences is assessed via distortion, which is formally defined as follows.

Definition 1.2 (Distortion). For a metric space $(U, d) \in \mathcal{M}$ and an ordinal profile $\succ_U \in \mathcal{P}(d)$, let $\mathcal{I}_{\succ_U}$ be the collection of instances of the ordinal (fair) k-center problem defined on points U and ordinal profile $\succ_U$. Let $\mathcal{A}(I \mid d)$ denote the solution returned by algorithm $\mathcal{A}$ on instance $I \in \mathcal{I}_{\succ_U}$ with underlying metric d. Let $S^*_{I,d}$ be an optimal solution to I when the underlying metric is d. The distortion of $\mathcal{A}$ is defined as:

$$\texttt{distortion}(\mathcal{A}) := \sup_{\substack{(U,d)\in\mathcal{M}\\ \succ_U\in\mathcal{P}(d)}} \sup_{I\in\mathcal{I}_{d,\succ_U}} \frac{\texttt{cost}(\mathcal{A}(I \mid d))}{\texttt{cost}(S^*_{I,d})}$$

Finally, our goal is to design algorithms with constant distortion guarantees using only near-linear (in k) many distance queries. Our key contribution is a 5-distortion algorithm that uses $O(k \log^2 k)$ queries; along the way, we give a 3-distortion algorithm using $O(k^2)$ queries. The full version of this manuscript and detailed proofs are available in arXiv [7].

2 OUR TECHNIQUES AND RESULTS

Our techniques build upon the maximum matching framework of [6, 9]. Starting from the "fair-oblivious" centers obtained by a Gonzalez-style greedy algorithm [3, 8], we project them—via a maximum matching framework—onto a set of centers that satisfies the fairness requirements, while controlling the additional cost incurred during this projection. We begin by characterizing the Gonzalez-style solution through the notion of *progressive cover*; we then introduce the *projection graph* that formalizes the matching-based projection to fair centers.

Progressive cover. Let $T = (t_1, \ldots, t_k)$ be an ordered set of "fair-oblivious" centers and let $T_\ell = (t_1, \ldots, t_\ell)$ denote its length-ℓ prefix. Let S^* be an optimal solution and $\mathbf{\Pi}^*$ be the partition of U induced by S^*. We say that T is a progressive γ-cover (w.r.t. S^*) if there exists some ℓ such that: (*i*) T_ℓ hits each cluster of $\mathbf{\Pi}^*$ at most once, and (*ii*) $\texttt{cost}(T_\ell) \le \gamma \cdot \texttt{cost}(S^*)$. We refer the largest such ℓ the critical index ℓ^*. A Gonzalez-style algorithm for ordinal k-center yields a set T that is a progressive 2-cover using $\frac{k(k-1)}{2}$ queries [3, Theorem 3.1] and a progressive 4-cover with $2k$-queries [3, Theorem 3.3].

Projection graph. Let H^ℓ_λ be the bipartite (ℓ, λ)-projection graph with left side T_ℓ and right side $\mathbb{G} = \{G_1, \ldots, G_k\}$; there is an edge from t_i to G_j iff $d(t_i, G_j) \le \lambda$. Let λ_ℓ be the minimum λ such that H^ℓ_λ admits a left-perfect matching (*i.e.*, every center in T_ℓ can be matched to a distinct group in $\mathbb{G}$). Given a left-perfect matching on $H^\ell_{\lambda_\ell}$, we construct a feasible solution S_ℓ as follows: for each matched pair (t_i, G_j), select the point in G_j closest to t_i. For any groups left unmatched, select an arbitrary point from that group. Let $Sol(\ell)$ be the cost of the feasible solution S_ℓ; by triangle inequality, it is at most $\texttt{cost}(T_\ell)+\lambda_\ell$. For the critical index ℓ^*, we have $\lambda_{\ell^*} \le \texttt{cost}(S^*)$ and $\texttt{cost}(T_{\ell^*}) \le \gamma \cdot \texttt{cost}(S^*)$, hence $Sol(\ell^*) \le (\gamma + 1) \cdot \texttt{cost}(S^*)$.

From k centers in T to the n points in $\mathbb{G}$, there are at most nk distinct distances; these values can be obtained by querying all nk center to point distances. In each iteration $\ell \in [k]$, since λ_{ℓ^*} must be one of these nk distances, we can identify λ^*_ℓ (*i.e.*, the optimal radius) and thus obtain 3 distortion (and 5 distortion resp.) by combining the 2-distortion (and 4-distortion resp.) algorithm of [3] for obtaining T with the matching frameworks of [6, 9]. This naturally leads to the following question: *can we achieve the same distortion guarantees with smaller number of queries to the cardinal information, preferably near-linear in k?*

Key results. To reduce the number of queries, our first idea is to shrink the search space of candidate distances from nk to k^2. For each center $t_i \in T$ and group $G_j \in \mathbb{G}$, we identify the nearest point in G_j to t_i using only the ordinal ranking of t_i. This results in at most k^2 distances to construct bipartite graph that facilitate us in finding the λ^*_ℓ for any $\ell \in [k]$. This yields a 3-distortion algorithm using $O(k^2)$ queries.

A possible way to further reduce the number of queries is to conduct binary search on ℓ over function $Sol(\ell) := \texttt{cost}(T_\ell) + \lambda_\ell$. However, this is not possible: $cost(T_\ell)$ is non-increasing and λ_ℓ is non-decreasing, their sum $Sol(\ell)$ is not necessarily monotone. To overcome this limitation, our first key contribution is introducing a predicate

$$P(\ell) \equiv (4\lambda_\ell \le cost(T_\ell)),$$

which *is monotone* and facilitates us to do binary search. We find the largest ℓ, denoted as L, such that $P(L)$ is true but $P(L + 1)$ is false. We further prove that $\min\{Sol(L), Sol(L + 1)\} \le 5\, cost(S^*)$, where S^* is the optimal solution, thus establishing the correctness of our binary-search procedure. Next, we bound the query complexities for evaluating $P(\ell)$ and λ_ℓ.

For computing λ_ℓ, we reduce the search space for λ_ℓ (with size k^2) by a factor of $\frac{3}{4}$, at each iteration. To accomplish this, at each iteration, we identify a *pivot point* in the search space such that at least $\frac{1}{4}$ of the candidate distances are no smaller than the pivot and at least $\frac{1}{4}$ are no greater than the pivot such that at least one part does not contain λ_ℓ. Finally, using a median-of-medians style subroutine, we find such a pivot; this results in query complexity $O(\ell \log^2 k)$. Notice that, for the complete algorithm, we only need to compute λ_ℓ for a fixed ℓ twice. Therefore, in the worst-case when $\ell = k$, the total query complexity of computing λ_ℓ is $O(k \log^2 k)$.

For evaluating $P(\ell)$, we apply its equivalent formulation $P(\ell) \equiv \left(\lambda_\ell \le \frac{cost(T_\ell)}{4}\right)$, which holds precisely when there exists a left-perfect matching when λ is set to $\frac{cost(T_\ell)}{4}$. Using this design, we obtain an overall query complexity of $O(k \log^2 k)$. Finally, we present our main result in the following theorem.

Theorem 2.1. *There exists a* 5*-distortion algorithm for the ordinal fair k-center problem that uses at most* $O(k \log^2 k)$ *queries.*

REFERENCES

[1] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence* 264 (2018), 27–51.

[2] Tiffany D. Barnes and Mirya R. Holman. 2020. Gender Quotas, Women's Representation, and Legislative Diversity. *The Journal of Politics* 82, 4 (2020), 1271–1286.

[3] Jakob Burkhardt, Ioannis Caragiannis, Karl Fehrs, Matteo Russo, Chris Schwiegelshohn, and Sudarshan Shyam. 2024. Low-distortion clustering with ordinal and limited cardinal information. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 9555–9563.

[4] Javier Cembrano and Golnoosh Shahkarami. 2025. Metric Distortion in Peer Selection. *arXiv preprint arXiv:2502.21084* (2025).

[5] Aurel Croissant and Hans-Joachim Lauth (Eds.). 2019. *Ethnic Quotas, Political Representation, and Equity in Asia Pacific.* Springer.

[6] Ameet Gadekar, Aristides Gionis, and Suhas Thejaswi. 2025. Fair clustering for data summarization: Improved approximation algorithms and complexity insights. In *Proceedings of the ACM on Web Conference 2025*. 4458–4469.

[7] Ameet Gadekar, Aristides Gionis, Suhas Thejaswi, and Sijing Tu. 2025. Fair Committee Selection under Ordinal Preferences and Limited Cardinal Information. *arXiv preprint arXiv:2512.24934* (2025).

[8] Teofilo F Gonzalez. 1985. Clustering to minimize the maximum intercluster distance. *Theoretical computer science* 38 (1985), 293–306.

[9] Matthew Jones, Huy Nguyen, and Thy Nguyen. 2020. Fair k-centers via maximum matching. In *International conference on machine learning*. PMLR, 4940–4949.

[10] Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. 2019. Fair k-center clustering for data summarization. In *International Conference on Machine Learning*. PMLR, 3448–3457.

[11] Ariel D Procaccia and Jeffrey S Rosenschein. 2006. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*. Springer, 317–331.

[12] Haripriya Pulyassary and Chaitanya Swamy. 2025. Constant-Factor Distortion Mechanisms for k-Committee Election. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 14062–14069.