Fair Clustering: Concepts, Methods, and Algorithms (Part I)

Ali Vakilian
Department of Computer Science, Virgina Tech
vakilian@vt.edu

1 Introduction

Clustering is a foundational problem in machine learning and data analysis, aiming to partition a dataset into subsets (clusters) such that data points within a cluster are similar according to a given objective. Classical clustering algorithms, such as k-means, k-median, and k-center, have found widespread use in applications ranging from recommendation systems and market segmentation to medical diagnostics and image analysis. However, these traditional objectives focus solely on optimizing notions of proximity or compactness and often overlook issues of *fairness*, particularly in how individuals from different demographic or social groups are treated.

In recent years, there has been a growing awareness that standard clustering procedures may exacerbate bias or reinforce structural disparities, such as forming clusters that disproportionately represent one demographic group while marginalizing others. This has motivated the development of *fair clustering* algorithms, which aim to incorporate fairness constraints alongside standard clustering objectives. Fairness in clustering can take many forms, including ensuring balanced group representation within clusters, minimizing worst-case costs across demographic groups (social fairness), or guaranteeing that similar individuals are treated similarly (individual fairness). These models bring together algorithmic design with ethical and societal considerations, making the study of fair clustering both technically rich and socially important.

In this survey, we provide an overview of algorithmic results in fair clustering, with an emphasis on theoretical models, approximation guarantees, and algorithmic techniques. Our focus is primarily on *centroid-based clustering* objectives, such as fair variants of k-center, k-median, and k-means, which have been the most extensively studied from an algorithmic perspective. We briefly mention other clustering frameworks where fairness has also been explored, including *spectral clustering* [Kleindessner et al., 2019b, Wang et al., 2023, Tonin et al., 2025] and *hierarchical clustering* [Ahmadian et al., 2020, Chhabra et al., 2021, Knittel et al., 2023a,b], though these are not the main focus of this survey.

It is worth emphasizing that this is by no means a comprehensive survey. The body of work in fair clustering is vast and growing, including numerous models, fairness criteria, algorithmic strategies, and empirical evaluations. Our goal is to present a curated overview of core algorithmic contributions that have shaped the field, particularly in terms of theoretical guarantees and complexity results. Moreover, we are not covering the empirical aspects of fair clustering in detail, such as the development of real-world benchmarks, practical deployment challenges, and critical evaluations of fairness definitions. These components are not only important but also essential for advancing the overall impact of fair clustering within the broader fields of algorithmic fairness and machine learning. Developing representative datasets and benchmarks is crucial for translating theoretical insights into effective practice, and deployment-focused studies often reveal practical trade-offs that are not captured by algorithmic models alone. For readers interested in these empirical and application-oriented directions, we refer to works such as [], which investigate the effectiveness of fair clustering on real datasets, examine benchmark design, and assess the interplay between fairness, accuracy, and interpretability in real-world settings. Well-known real-world examples, such as the ProPublica analysis of recidivism risk scores in classification [Angwin et al., 2022], have been instrumental in highlighting both the challenges and societal impact of fairness in algorithmic decision making. Similarly, compelling case studies and practical applications in fair clustering have the potential to capture attention and clearly demonstrate the real-world value and necessity of fairness in this area.

Roadmap of the survey. Fair clustering methods can be broadly categorized according to the type of fairness they seek to enforce, reflecting different philosophical and practical interpretations of what it means for a clustering to be *"fair"* [Dwork et al., 2012, Hardt et al., 2016, Barocas et al., 2017, Mitchell et al., 2021]. These categories distinguish whether fairness is defined at the level of individuals, groups, or the overall distribution of clustering outcomes, and each perspective introduces unique algorithmic challenges and modeling considerations.

2 Group Fairness Notions in Clustering

A central line of work in fair clustering focuses on enforcing *group fairness*, where individuals are partitioned into protected demographic groups (such as by race, gender, or age), and the goal is to ensure that the clustering process does not result in disproportionate or disparate treatment of any group. Group fairness has attracted significant interest due to its close connection to real-world equity concerns and its potential to reduce bias in downstream decision-making applications. The key idea is to ensure that no group is systematically underrepresented, overrepresented, or disproportionately burdened by the clustering output.

However, there is no single canonical definition of group fairness in clustering. Instead, several formalizations have been proposed, each capturing a different intuition about what it means for a clustering to be fair. These include notions such as *balance within clusters*, where the demographic composition of each cluster should resemble that of the whole dataset; *fair representation within centers*, where the selected cluster representatives (e.g., medoids or means) should be demographically diverse; and *cost-based fairness* such as min-max or *socially fair clustering*, which aims to equalize clustering costs across groups.

In the following subsections, we survey key algorithmic developments under each of these group fairness paradigms.

2.1 Fair representation within clusters

This notion of group fairness in clustering aims to ensure that each cluster reflects the demographic composition of the dataset, particularly across protected groups such as gender or ethnicity. Chierichetti et al. [2017] pioneered the algorithmic study of fair clustering by introducing the notion of *balance*, a quantitative measure of group parity within a subset.

In their seminal work, Chierichetti et al. [2017] introduced the notion of fair clustering with two protected groups. In this setting, we are given a set of n points P in a metric space (X, d), where

Figure 1: **Left (U):** An unbalanced clustering, where the selected cluster is dominated by individuals from one group (green), with only a single member from the other group, resulting in unequal representation in the advantaged cluster. **Right (F):** A more balanced clustering.

each point belongs to either the *red* or the *blue* group. For any subset of points $S \subseteq P$, let $r(S)$ and $b(S)$ denote the number of red and blue points in S, respectively. The balance is defined as $\mathrm{balance}(S) = \min\left(\frac{r(S)}{b(S)}, \frac{b(S)}{r(S)}\right)$, which attains its maximum value of 1 when the groups are equally represented. Given the point set P, a number of clusters k, and a required minimum balance $t \in (0, 1]$, the fair clustering problem is to partition P into clusters $\mathcal{C} = \{C_1, \ldots, C_k\}$ such that each cluster C_i satisfies $\mathrm{balance}(C_i) \geq t$, a property referred to as t-balance, while minimizing a clustering objective such as k-center or k-median.

Chierichetti et al. [2017] showed that enforcing exact fairness constraints in clustering is **NP**-hard, and then develop a two-stage approximation framework based on the concept of *fairlets* which are small, t-balanced subsets that act as "atomic" units in the subsequent clustering process. The method first decomposes the dataset into a collection of fairlets and then applies standard clustering algorithms on their representatives. For the case of $t = 1$, the fairlet decomposition reduces to a minimum-cost perfect matching problem. Their framework yields approximation algorithms when the balance parameter t is a rational number: For example, they provide a 4-approximation for the fair k-center problem, and show $O(t)$-approximation for the fair k-median problem.

This two-step method preserves fairness guarantees and modularity which allows existing clustering algorithms to be adapted with minimal changes. However, the construction of fairlets may become computationally expensive, particularly as the number of groups increases.

Next, we consider a generalization of the fair clustering problem to the setting with multiple groups. Specifically, we are given a set of n points P in a metric space (X, d), where each point may belong to one or more of ℓ (possibly overlapping) groups $P_1, \ldots, P_\ell$, and $P = \bigcup_{i \in [\ell]} P_i$.

2.1.1 Extending to multiple groups: (I) absolute balance.

Generalizing the fair representation notion of Chierichetti et al. [2017], Bera et al. [2019] introduced a more general model of fairness in clustering using parameterized group constraints applicable to the setting with more than two groups. Each protected group j is assigned lower and upper fraction bounds, α_j and β_j, to enforce representation constraints in every cluster. Formally, a clustering $\mathcal{C} = \{C_1, \cdots, C_k\}$ is (α, β)-*fair* if for every cluster $C \in \mathcal{C}$ and every group $j \in [\ell]$, $\alpha_j \cdot |C| \leq |C \cap P_j| \leq \beta_j \cdot |C|$.

Bera et al. [2019] present a *post-processing algorithm* that converts any vanilla clustering (for any ℓ_p objective, such as k-means, k-median, or k-center) into one that satisfies these fairness constraints, with an additive violation of at most $4\Delta + 3$ individuals per group per cluster (where Δ is the maximum number of groups any point belongs to), while increasing the clustering cost by at most a constant factor. Independently, Bercea et al. [2019] obtained similar guarantees by allowing small additive violations in group representation. While both approaches seek to balance fairness and clustering quality, the algorithm of Bera et al. [2019] provides stronger guarantees and supports the more general setting with overlapping group memberships.

For fair k-center clustering, Harb and Lam [2020] proposed a scalable algorithm that achieves a 5-approximation guarantee, with additive fairness violations matching those in Bera et al. [2019]. Notably, although this algorithm allows for worst-case additive violations of fairness constraints, the authors show that the expected additive deviation from the specified α_i and β_i bounds is zero[1].

By designing coresets for fair clustering, Bandyapadhyay et al. [2021] presented algorithms that compute a constant-factor approximation for fair k-median and k-means, running in time $(k\Delta)^{O(k\Delta)}\text{poly}(n)$, and crucially, these algorithms do not violate the fairness constraints. In the special case of no overlap between groups, the runtime improves to $(k\ell)^{O(k\ell)}\text{poly}(n)$. Previously, coresets for fair representation clustering were also studied in [Schmidt et al., 2019, Huang et al., 2019].

Prior to [Bera et al., 2019], this notion of fairness was studied for (1) the k-center problem [Rösner and Schmidt, 2018], and (2) k-clustering with an ℓ_p-objective on balanced instances, which are instances with equal group sizes and $\alpha_j = \beta_j = 1$ for all $j \in [\ell]$ [Böhm et al., 2021]. In these special cases, the fair clustering problem admits a constant-factor approximation. In another line of research, a special case of this fairness notion, known as *restricted dominance*, imposes only upper bounds on the fraction of each group within clusters. For this setting, Ahmadian et al. [2019] developed a bicriteria approximation algorithm that provides fairness guarantees with bounded additive violations

Building on [Bera et al., 2019, Backurs et al., 2019], Dai et al. [2022] designed a "purely multiplicative" $O(\log k)$-approximation for the k-median objective (i.e., without additive violations) in the setting with multiple groups, under the (α, β) fairness notion of [Bera et al., 2019]. Their algorithm runs in time $n^{O(\ell)}$, where ℓ is the number of groups. It was subsequently shown by Bandyapadhyay et al. [2024] that obtaining an approximation factor of $o(\log k)$ for fair k-median in *polynomial time* is as hard as breaking the long-standing $\Omega(\log k)$ barrier for the well-studied *soft uniform capacitated k-median* problem. The approach of Dai et al. [2022] relies on probabilistically approximating arbitrary metrics by tree metrics [Bartal, 1998, Fakcharoenphol et al., 2003]. Notably, in the context of fair k-median, this idea of approximating the input metric with a distribution over dominating trees was also used by Backurs et al. [2019] in their *near-linear time* approximation algorithm for fair representation clustering with two groups.

Question 1. *Can the $O(\log k)$-approximation barrier for fair k-median under multi-group (α, β) fairness constraints be improved in polynomial time, or is this limitation fundamental?*

Question 2. *Is it possible to design a near-linear time approximation algorithm for fair k-clustering with general ℓ_p objectives under multi-group (α, β) fairness constraints? In particular, can we obtain algorithms that are both efficient and provide strong approximation guarantees for settings beyond two groups and objectives beyond k-median?*

[1] That is, the group representation bounds are satisfied on average, as opposed to expected zero violations.

2.1.2 Extending to multiple groups: (II) pairwise

In a different extension of the fair representation within clusters to multiple groups, Bandyapadhyay et al. [2024, 2025] explicitly studied the notion of *pairwise* fairness which given an integer parameter t, requires to output a clustering $\mathcal{C} = \{C_1, \cdots, C_k\}$ such that for every cluster $C \in \mathcal{C}$, and every pair of groups $i, j \in [\ell]$, $|C \cap P_i|/t \leq |C \cap P_j| \leq t \cdot |C \cap P_i|$.

For k-center, via an LP-based rounding scheme, Bandyapadhyay et al. [2025] show an $O(1)$-approximation for two settings: (i) there are two groups in the input and t is an integer, and (ii) when $t = 1$. On the hardness side, the authors prove an integrality gap of $\Omega(k)$ for their proposed LP relaxation when t is not an integer even with two groups.

For k-median, the problem turns out to be more difficult. Bandyapadhyay et al. [2024] provides $O(k^2\ell t)$-approximation for t-pairwise fair k-clustering when there are ℓ disjoint groups in the input. On the negative side, they show it is NP-hard to approximate the problem within a factor better than $n^{1-\varepsilon}$, for every constant $\varepsilon > 0$, when there are *overlapping groups*, even for $k = 2$.

2.1.3 Further remarks

Fair clustering with Sum-of-Radii objective. This notion has been popular and well-studied for clustering with other objective function such as sum-of-radii, where the the goal is to minimize the sum of radii of the constructed clusters [Carta et al., 2024, Nezhad et al., 2025].

Fairness under bounded cost. Another research direction focuses on minimizing unfairness given an upper bound on the clustering cost, as proposed by Esmaeili et al. [2021]. Instead of minimizing clustering cost subject to strict fairness constraints, this approach sets an upper limit on the clustering cost and shifts the objective to minimizing unfairness. In this setting, each point belongs to a group, and clusters are required to approximately reflect the population proportions of each group. Formally, unfairness for group j is defined as the minimum proportional relaxation δ_j such that, in every cluster, the group's proportion lies within the relaxed bounds $[\alpha_j - \delta_j, \beta_j + \delta_j]$. The paper considers three natural objectives: *group-utilitarian* (minimize the sum of all group violations), *group-egalitarian* (minimize the maximum group violation), and *group-leximin* (minimize the worst, then the second-worst, and so on, violations). The optimization problem can be formally stated as minimizing one of these *unfairness objectives*, subject to the constraint that the clustering cost does not exceed a given threshold U. For the first two fairness objectives, the authors provide bicriteria approximation algorithms that achieve bounded violations of the fairness constraints, and ensure that the objective value is within a small additive error of the upper bound U.

Fair labeled clustering. Clustering is a widely used preprocessing step in many decision-making pipelines. From a fairness perspective, it is often sufficient to ensure equity in the outcomes or labels assigned to clusters in subsequent steps, rather than enforcing group balance within the clusters themselves. Motivated by these applications, **?** introduced the fair labeled clustering framework, where fair representation constraints are imposed on the set of points that receive each outcome label. Specifically, after clusters are formed, the goal is to ensure that, for each label, the aggregate population assigned that label satisfies the desired group representation, even if individual clusters do not. This approach is less restrictive than traditional fair clustering and often results in higher-quality clusterings, while still mitigating disparate impact in the outcomes.

The fair labeled clustering problem has two main variants. In the first, *labeled clustering with assigned labels (LCAL)*, each cluster is given a predetermined outcome label (for example, based on the cluster center). In the second, *labeled clustering with unassigned labels (LCUL)*, labels can be assigned to clusters in a way that further improves fairness or efficiency. In both cases, the goal is to minimize clustering cost while satisfying proportional group representation constraints for each label. Notably, LCAL can be solved in polynomial time for any fixed number of labels, in contrast to the traditional fair clustering problem, which is **NP**-hard (even not known whether it is possible to approximate within a factor better than $\Omega(\log k)$). The authors develop efficient algorithms based on network flow for both settings, as well as randomized rounding for LCUL to ensure fairness constraints hold in expectation.

Distance to fairness and consensus. Chakraborty et al. [2025] introduced the problem of finding the closest fair clustering to a given clustering in the case of binary groups, and devised a constant-factor approximation algorithm for it. They also proposed the notion of fair consensus clustering, where, given a collection of clusterings $\mathcal{D}_1, \ldots, \mathcal{D}_m$ over the same dataset, the goal is to find a fair clustering C^* that minimizes the overall distance to the input clusterings. The distance is defined as the ℓ_p-norm of the vector of distances from C^* to each $\mathcal{D}_i$, where the distance between two clusterings is the number of pairs of points that are clustered together in one clustering but not in the other. For this consensus problem, they show that their approach for closest fair clustering yields a constant-factor approximation in the two-group setting.

Question 3. *Is it possible to generalize the results of Chakraborty et al. [2025] for closest fair clustering and fair consensus clustering to settings with more than two groups?*

In large-scale computational models. In the context of fair clustering, with fair representation within clusters, under massive data processing models, streaming algorithms have received increasing attention due to the need to process large-scale and rapidly arriving data. A key challenge in this setting is to ensure fairness, typically defined via demographic constraints or group-based balance, while maintaining computational and memory efficiency. In the standard insertion-only streaming model, several works have explored the use of *coresets* for fair k-means and k-median clustering [Schmidt et al., 2019, Huang et al., 2019, Bandyapadhyay et al., 2021], providing algorithms that maintain compact summaries (coresets) of the input while approximately preserving both clustering cost and fairness constraints. These approaches allow for efficient post-processing to recover fair clusterings and have established bounds on coreset sizes and approximation guarantees. However, adapting fairness to more dynamic models like the *sliding window model*, where only the most recent elements are considered, remains more challenging. The recent work of Cohen-Addad et al. [2025] initiates the study of fair clustering in this setting, providing the first algorithms for fair k-center clustering in sliding windows with provable approximation and space guarantees.

2.1.4 Fair representation within centers (Centroid-based clustering)

Another important fairness notion in clustering is *fair representation among centers*, where the demographic composition of the selected cluster centers themselves must reflect a desired distribution. This idea is particularly relevant in centroid-based clustering objectives such as k-center, k-median, and k-means, especially when the centers are intended to serve as a representative summary of the dataset (e.g., in applications like data summarization or decision-making).

Kleindessner et al. [2019a] introduced a formulation where the dataset is partitioned into ℓ demographic groups, and exactly k_i centers must be chosen from group i, for some input parameters $k_1, \ldots, k_\ell$. While this ensures strict demographic proportionality, it may significantly degrade clustering quality. To address this, [Hotegni et al., 2023, Nguyen et al., 2022] proposed a relaxation called *fair range clustering*, which requires the number of selected centers from each group i to fall within a specified interval $[\alpha_i, \beta_i]$. This flexible model captures a range of fairness policies while allowing for improved clustering performance. They present constant-factor approximation algorithms for fair range k-clustering under ℓ_p-norm objectives, including k-center, k-median, and k-means.

Overall, fair representation within centers is crucial in applications like image search, recourse design, or any setting where the selected representatives are presented directly to users or downstream decision-making systems.

Further remarks. Clustering with constraints on the *set of centers* that form an independent set of a matroid over the candidate facilities is a well-studied abstraction that subsumes this notion of fairness. This line of work has yielded constant-factor approximations through LP relaxations with matroid polytopes, swap/pipage rounding, and matroid-aware local search, across median/means/center-style objectives, and has been used extensively in the fair-clustering literature [Hajiaghayi et al., 2010, Krishnaswamy et al., 2011, Charikar and Li, 2012, Swamy, 2016, Chen et al., 2016, Krishnaswamy et al., 2018, Jones et al., 2020, Chiplunkar et al., 2020]. In particular, the notion of fairness with exact per-group quotas is a direct special case (a partition-matroid feasibility on centers), placing it squarely within this framework; by contrast, the *fair range* variant relaxes quotas to intervals and is not itself a matroid (since feasibility is not downward-closed), though it admits reductions that leverage matroid structure via extendable-set relaxations in [El Halabi et al., 2020, Hotegni et al., 2023].

2.1.5 Min-Max fairness (aka socially fair clustering)

Socially fair clustering addresses disparities in clustering cost across different demographic groups. Rather than requiring fair representation within each cluster, the objective here is to minimize the worst-case (maximum) clustering cost incurred by any group. Early work by Abbasi et al. [2021] and Ghadiri et al. [2021] introduced this formulation for k-median and k-means objectives, providing $O(\ell)$-approximation algorithms, where ℓ is the number of protected groups. However, their results relied on natural LP relaxations that were shown to have an integrality gap of $\Omega(\ell)$, limiting the scope for improved approximations.

To address this limitation, [Makarychev and Vakilian, 2021] introduced a strengthened LP relaxation for socially fair ℓ_p-clustering that adds locality constraints tailored to the structure of the metric space and the weight distribution. Their algorithm achieves an $\tilde{O}(\log \ell / \log\log \ell)$-approximation for general p, which is tight under standard complexity assumptions—matching a hardness lower bound for even unweighted k-median on uniform metrics Bhattacharya et al. [2014]. This result represents the best-known approximation for socially fair k-median and k-means (for $p = 1$ and $p = 2$), and generalizes to arbitrary weights over groups. Moreover, [Makarychev and Vakilian, 2021] extend their techniques to the weighted case, where each group j is represented by a weight function $w_j : L \to \mathbb{R}_+$ over the set of data points, and the goal is to minimize $\max_j \sum_{v \in L} w_j(v) \cdot d(v, C)^p$ for a center set C. This framework captures applications where group importance varies across the domain, and it broadens applicability to non-disjoint and even overlapping group memberships.

Finally, their algorithm supports both multiplicative and bicriteria approximation guarantees. The bicriteria version relaxes the number of centers from k to $k/(1-\gamma)$ and achieves a cost approximation of $O(1/\gamma)$ times the optimum. Overall, this line of work solidifies min-max fairness as a practically and theoretically robust approach to equitable clustering across heterogeneous populations.

Chlamtác et al. [2022] extended the socially fair clustering framework to a more general family of objectives called (p, q)-fair clustering, where the goal is to minimize the ℓ_q norm across groups of the ℓ_p clustering cost incurred by each group. This formulation captures and generalizes a wide range of clustering problems: classical clustering ($q = p$), robust clustering ($q = 1$) Anthony et al. [2010], and socially fair clustering ($q = \infty$). The (p, q) objective provides a flexible way to interpolate between average-case fairness and worst-case fairness, allowing practitioners to tune the fairness strictness by adjusting q relative to p.

The authors developed approximation algorithms based on novel convex programming relaxations. For the case $p \leq q$, they achieved an $O((q/\ln(1 + q/p))^{1/p})$-approximation using randomized rounding and careful handling of integrality gaps through non-standard constraints. For the case $p \geq q$, they presented an $O(k^{(p-q)/(2pq)})$-approximation, which nearly matches their derived hardness bound of $k^{\Omega((p-q)/(pq))}$ under standard assumptions, such as the hardness of the Min-s-Union problem. These results nearly close the gap between algorithmic upper bounds and known lower bounds for various p, q regimes. Overall, the (p, q)-fair clustering model provides a powerful framework for balancing fairness and efficiency in clustering, accommodating a range of practical fairness requirements beyond strict max-min fairness.

Goyal and Jaiswal [2023] provided *tight fixed-parameter tractable (FPT)* $(3^z + \varepsilon)$-approximation algorithms for socially fair k-median ($z = 1$) and k-means ($z = 2$), running in time $(k/\varepsilon)^{O(k)} \cdot n^{O(1)}$, and proved these bounds to be optimal under the *Gap-ETH* assumption.

The min-max fairness framework has since been influential in formalizing fairness in clustering beyond compositional constraints, addressing real-world scenarios that demand cost equity over demographic parity.

Conclusion. Fair clustering is a vibrant area at the intersection of algorithm design and social responsibility. This survey reviewed core fairness definitions and highlighted key algorithmic techniques, with an emphasis on group fairness notions. In future parts of this survey, we will cover additional notions of fairness for clustering, in particular, those related to individual fainress.

References

Mohsen Abbasi, Aditya Bhaskara, and Suresh Venkatasubramanian. Fair clustering via equitable group representations. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 504–514, 2021.

Sara Ahmadian, Alessandro Epasto, Ravi Kumar, and Mohammad Mahdian. Clustering without over-representation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 267–275, 2019.

Sara Ahmadian, Alessandro Epasto, Marina Knittel, Ravi Kumar, Mohammad Mahdian, Benjamin Moseley, Philip Pham, Sergei Vassilvitskii, and Yuyan Wang. Fair hierarchical clustering. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias. In *Ethics of data and analytics*, pages 254–264. 2022.

Barbara Anthony, Vineet Goyal, Anupam Gupta, and Viswanath Nagarajan. A plant location guide for the unsure: Approximation algorithms for min-max location problems. *Mathematics of Operations Research*, 35(1):79–101, 2010.

Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, and Tal Wagner. Scalable fair clustering. In *International conference on machine learning*, pages 405–413. PMLR, 2019.

Sayan Bandyapadhyay, Fedor V. Fomin, and Kirill Simonov. On coresets for fair clustering in metric and euclidean spaces and their applications. In *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 23:1–23:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

Sayan Bandyapadhyay, Eden Chlamtáč, Zachary Friggstad, Mahya Jamshidian, Yury Makarychev, and Ali Vakilian. A polynomial-time approximation for pairwise fair k-median clustering. *arXiv preprint arXiv:2405.10378*, 2024.

Sayan Bandyapadhyay, Tianzhi Chen, Zachary Friggstad, and Mahya Jamshidian. A constant-factor approximation for pairwise fair k-center clustering. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 43–57. Springer, 2025.

Solon Barocas, Moritz Hardt, and Arvind Narayanan. Fairness in machine learning. *arXiv preprint arXiv:1908.09635*, 2017.

Yair Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 161–168, 1998.

Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. Fair algorithms for clustering. *Advances in Neural Information Processing Systems*, 32, 2019.

Ioana Oriana Bercea, Martin Groß, Samir Khuller, Aounon Kumar, Clemens Rösner, Daniel R. Schmidt, and Melanie Schmidt. On the cost of essentially fair clusterings. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA*, volume 145 of *LIPIcs*, pages 18:1–18:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

Sayan Bhattacharya, Parinya Chalermsook, Kurt Mehlhorn, and Adrian Neumann. New approximability results for the robust k-median problem. In *Scandinavian Workshop on Algorithm Theory*, pages 50–61. Springer, 2014.

Matteo Böhm, Adriano Fazzone, Stefano Leonardi, Cristina Menghini, and Chris Schwiegelshohn. Algorithms for fair k-clustering with multiple protected attributes. *Oper. Res. Lett.*, 49(5):787–789, 2021.

Lena Carta, Lukas Drexler, Annika Hennes, Clemens Rösner, and Melanie Schmidt. Fpt approximations for fair k-min-sum-radii. *arXiv preprint arXiv:2410.00598*, 2024.

Diptarka Chakraborty, Kushagra Chatterjee, Debarati Das, Tien Long Nguyen, and Romina Nobahari. Towards fair representation: Clustering and consensus. *arXiv preprint arXiv:2506.08673*, 2025.

Moses Charikar and Shi Li. A dependent lp-rounding approach for the k-median problem. In *International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 194–205, 2012.

Danny Z Chen, Jian Li, Hongyu Liang, and Haitao Wang. Matroid and knapsack center problems. *Algorithmica*, 75(1):27–52, 2016.

Anshuman Chhabra, Adish Singla, and Prasant Mohapatra. Fair clustering using antidote data. In Jessica Schrouff, Awa Dieng, Miriam Rateike, Kweku Kwegyir-Aggrey, and Golnoosh Farnadi, editors, *Algorithmic Fairness through the Lens of Causality and Robustness Workshop, AFCR 2021, virtual, December 13, 2021*, volume 171 of *Proceedings of Machine Learning Research*, pages 19–39. PMLR, 2021.

Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5029–5037, 2017.

Ashish Chiplunkar, Sagar Kale, and Sivaramakrishnan Natarajan Ramamoorthy. How to solve fair k-center in massive data models. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 1877–1886, 2020.

Eden Chlamtác, Yury Makarychev, and Ali Vakilian. Approximating fair clustering with cascaded norm objectives. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 2664–2683. SIAM, 2022.

Vincent Cohen-Addad, Shaofeng H-C Jiang, Qiaoyuan Yang, Yubo Zhang, and Samson Zhou. Fair clustering in the sliding window model. In *The Thirteenth International Conference on Learning Representations*, 2025.

Zhen Dai, Yury Makarychev, and Ali Vakilian. Fair representation clustering with several protected classes. In *FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, June 21 - 24, 2022*, pages 814–823. ACM, 2022.

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)*, pages 214–226, 2012.

Marwa El Halabi, Slobodan Mitrović, Ashkan Norouzi-Fard, Jakab Tardos, and Jakub M Tarnawski. Fairness in streaming submodular maximization: Algorithms and hardness. *Advances in Neural Information Processing Systems (NeurIPS)*, 33:13609–13622, 2020.

Seyed A. Esmaeili, Brian Brubach, Aravind Srinivasan, and John Dickerson. Fair clustering under a bounded cost. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 14345–14357, 2021.

Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 448–455, 2003.

Mehrdad Ghadiri, Samira Samadi, and Santosh Vempala. Socially fair k-means clustering. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 438–448, 2021.

Dishant Goyal and Ragesh Jaiswal. Tight FPT approximation for socially fair clustering. *Inf. Process. Lett.*, 182:106383, 2023.

MohammadTaghi Hajiaghayi, Rohit Khandekar, and Guy Kortsarz. Budgeted red-blue median and its generalizations. In *Proceedings of the European Symposium on Algorithms (ESA)*, pages 314–325, 2010.

Elfarouk Harb and Ho Shan Lam. KFC: A scalable approximation algorithm for k-center fair clustering. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Moritz Hardt, Eric Price, and Nathan Srebro. Equality of opportunity in supervised learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3315–3323, 2016.

Sedjro Salomon Hotegni, Sepideh Mahabadi, and Ali Vakilian. Approximation algorithms for fair range clustering. In *International Conference on Machine Learning*, pages 13270–13284. PMLR, 2023.

Lingxiao Huang, Shaofeng H.-C. Jiang, and Nisheeth K. Vishnoi. Coresets for clustering with fairness constraints. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 7587–7598, 2019.

Matthew Jones, Huy Nguyen, and Thy Nguyen. Fair k-centers via maximum matching. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 4940–4949, 2020.

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

Matthäus Kleindessner, Samira Samadi, Pranjal Awasthi, and Jamie Morgenstern. Guarantees for spectral clustering with fairness constraints. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 3458–3467. PMLR, 2019b.

Marina Knittel, Max Springer, John P. Dickerson, and MohammadTaghi Hajiaghayi. Generalized reductions: Making any hierarchical clustering fair and balanced with low cost. In *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, pages 17218–17242. PMLR, 2023a.

Marina Knittel, Max Springer, John P. Dickerson, and MohammadTaghi Hajiaghayi. Fair, polylog-approximate low-cost hierarchical clustering. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine, editors, *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023b.

Ravishankar Krishnaswamy, Amit Kumar, Viswanath Nagarajan, Yogish Sabharwal, and Barna Saha. The matroid median problem. In *Proceedings of the Symposium on Discrete Algorithms (SODA)*, pages 1117–1130, 2011.

Ravishankar Krishnaswamy, Shi Li, and Sai Sandeep. Constant approximation for k-median and k-means with outliers via iterative rounding. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 646–659, 2018.

Yury Makarychev and Ali Vakilian. Approximation algorithms for socially fair clustering. In *Conference on Learning Theory*, pages 3246–3264. PMLR, 2021.

Shira Mitchell, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum. Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, 8: 141–163, 2021.

Sina Bagheri Nezhad, Sayan Bandyapadhyay, and Tianzhi Chen. Polynomial-time constant-approximation for fair sum-of-radii clustering. *arXiv preprint arXiv:2504.14683*, 2025.

Huy Lê Nguyen, Thy Nguyen, and Matthew Jones. Fair range k-center. *arXiv preprint arXiv:2207.11337*, 2022.

Clemens Rösner and Melanie Schmidt. Privacy preserving clustering with constraints. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, pages 96–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2018.

Melanie Schmidt, Chris Schwiegelshohn, and Christian Sohler. Fair coresets and streaming algorithms for fair k-means. In Evripidis Bampis and Nicole Megow, editors, *Approximation and Online Algorithms - 17th International Workshop, WAOA 2019, Munich, Germany, September 12-13, 2019, Revised Selected Papers*, volume 11926 of *Lecture Notes in Computer Science*, pages 232–251. Springer, 2019.

Chaitanya Swamy. Improved approximation algorithms for matroid and knapsack median problems and applications. *ACM Transactions on Algorithms (TALG)*, 12(4):1–22, 2016.

Francesco Tonin, Alex Lambert, Johan AK Suykens, and Volkan Cevher. Accelerating spectral clustering under fairness constraints. *arXiv preprint arXiv:2506.08143*, 2025.

Ji Wang, Ding Lu, Ian Davidson, and Zhaojun Bai. Scalable spectral clustering with group fairness constraints. In Francisco J. R. Ruiz, Jennifer G. Dy, and Jan-Willem van de Meent, editors, *International Conference on Artificial Intelligence and Statistics, 25-27 April 2023, Palau de Congressos, Valencia, Spain*, volume 206 of *Proceedings of Machine Learning Research*, pages 6613–6629. PMLR, 2023.