Proportionally Fair Clustering Revisited

Evi Micha
University of Toronto, Canada
emicha@cs.toronto.edu

Nisarg Shah
University of Toronto, Canada
nisarg@cs.toronto.edu

─── **Abstract** ───────────────────────────────

In this work, we study fairness in centroid clustering. In this problem, k cluster centers must be placed given n points in a metric space, and the cost to each point is its distance to the nearest cluster center. Recent work of Chen et al. [8] introduces the notion of a *proportionally fair* clustering, in which no group of at least n/k points can find a new cluster center which provides lower cost to each member of the group. They propose a *greedy capture* algorithm which provides a $1 + \sqrt{2}$ approximation of proportional fairness for any metric space, and derive generalization bounds for learning proportionally fair clustering from samples in the case where a cluster center can only be placed at one of finitely many given locations in the metric space.

We focus on the case where cluster centers can be placed anywhere in the (usually infinite) metric space. In case of the L^2 distance metric over $\mathbb{R}^t$, we show that the approximation ratio of greedy capture improves to 2. We also show that this is due to a special property of the L^2 distance; for the L^1 and L^∞ distances, the approximation ratio remains $1 + \sqrt{2}$. We provide universal lower bounds which apply to all algorithms.

We also consider metric spaces defined on graphs. For trees, we show that an exact proportionally fair clustering always exists and provide an efficient algorithm to find one. The corresponding question for general graph remains an interesting open question.

Finally, we show that for the L^2 distance, checking whether a proportionally fair clustering exists and implementing greedy capture over an infinite metric space are NP-hard problems, but (approximately) solvable in special cases. We also derive generalization bounds which show that an approximately proportionally fair clustering for a large number of points can be learned from a small number of samples. Our work advances the understanding of proportional fairness in clustering, and points out many avenues for future work.

1 Introduction

Machine learning algorithms are increasingly being used for decision making in applications where they affect human lives; popular examples include resume screening, evaluation of loan applications, bail decisions, etc. [27]. This has led to growing concern as to whether these algorithms, which may view humans as "data points", treat them fairly [3, 25]. Consequently, research on designing fair machine learning algorithms is proliferating [24, 22].

Much of this literature focuses on fairness in classification [36, 35, 19], but the study of fairness in other settings such as regression [1] and clustering [8] is also on the rise. In this paper, we focus on fairness in clustering, specifically, in centroid clustering. In this problem, we are given a set $\mathcal{N}$ of data points in a metric space, and a set $\mathcal{M}$ of possible locations for cluster centers in the same metric space. Given $k \in \mathbb{N}$, the task is to select a set $X \subseteq \mathcal{M}$ consisting of $|X| = k$ cluster centers and assign each data point to a cluster center – usually the closest – with the goal that data points are close to their assigned cluster centers. Clustering has diverse applications in market research, pattern recognition, data analysis, image segmentation, and facility location. In applications like image segmentation or market research, often the goal is to simply identify different clusters of points among the data. However, in facility location [14, 21, 15], where data points may represent locations of houses in the neighborhood and cluster centers may represent locations where public facilities (such as parks) will be built, it is of paramount importance that the facilities be distributed to fairly serve the population.

Adapting an example given by Chen et al. [8], imagine that there is a dense urban area with a population of 10,000, and far from it, there are 10 small communities with a population of 100 each. The communities are closer to each other compared to how far they are from the urban area, but still well distinguished. With $k = 11$, a standard clustering algorithm such as k-means would identify the urban area as one cluster, and each small community as one cluster. However, building just one park that serves 10,000 people in the urban area, while each community of 100 people gets its own park violates the principle of *equal entitlement* [23]; this principle would suggest that when allocating 11 parks among a total of 11,000 people, the urban area consisting of 10,000 people should be allocated their proportional share of 10 parks, and one park should serve the 10 smaller communities consisting of 1,000 people altogether.

This notion of what a group deserves – *group fairness* – has been extensively studied in machine learning, and a variety of definitions have been proposed [7, 16, 24, 19, 32]. Borrowing from a long line of literature on fair resource allocation [34, 30, 12, 10], Chen et al. [8] proposed a novel definition of fairness in clustering that perfectly fits our motivation. Given a metric d over a set $\mathcal{N}$ of n points and a set $\mathcal{M}$ of feasible cluster centers, they say that a k-clustering X satisfies *proportional fairness* if there is no group of points $S \subseteq \mathcal{N}$ with $|S| \geq n/k$ and a new cluster location $y \in \mathcal{M}$ such that $d(i, y) < \min_{x \in X} d(i, x)$ for each member $i \in S$. Most fairness definitions in machine learning protect groups of individuals that are pre-defined based on certain *protected attributes* [4, 5, 9, 28] or that are sufficiently large [20]. In contrast, proportional fairness guarantees fairness to arbitrarily defined groups of all possible sizes. This may be helpful given recent observations that protecting groups defined based on individual attributes may allow an algorithm to circumvent fairness [20], or that information about which groups to protect may not be known in advance [17]. For references to other related work, we direct the reader to the work of Chen et al. [8].

1.1 Our Contribution

We build upon the work of Chen et al. [8]. While their work considers the metric d and the set of feasible cluster center locations $\mathcal{M}$ to be arbitrary (and $|\mathcal{M}|$ to be typically finite), we focus on the case where the metric consists of usual distance functions such as L^1, L^2, or L^∞ over $\mathbb{R}^t$, and cluster centers can be placed anywhere in the infinite metric space (i.e. $\mathcal{M} = \mathbb{R}^t$). While this change is seemingly simple, the infinite cardinality of $\mathcal{M}$ requires new algorithmic tools and generalization bounds, which we provide in this work. In some cases, we show that this in fact allows us to provide stronger approximation guarantees.

In Section 3, we analyze the greedy capture algorithm introduced by Chen et al. in the case where $d \in \{L^1, L^2, L^\infty\}$ and $\mathcal{M} = \mathbb{R}^t$. Chen et al. show that the algorithm provides $1 + \sqrt{2} \approx 2.414$ approximation to proportional fairness for all metric spaces. We show that for $d = L^2$ and $\mathcal{M} = \mathbb{R}^t$, it actually provides a better 2-approximation. We prove this via a refinement of the result of Chen et al.: we express the approximation ratio obtained by the algorithm in terms of a new characteristic of the metric that we term *Apollonius radius*, and show that this radius is small for the L^2 distance, allowing us to achieve a better approximation ratio. However, we show that for L^1 and L^∞, the approximation ratio of greedy capture is no better than $1 + \sqrt{2}$.

In Section 4, we provide universal lower bounds which apply to all algorithms. Specifically, we show that for $d = L^2$ and $\mathcal{M} = \mathbb{R}^t$, no algorithm achieves better than $2/\sqrt{3} \approx 1.155$ approximation ratio, whereas for $d \in \{L^1, L^\infty\}$ and $\mathcal{M} = \mathbb{R}^t$, we get a lower bound of 1.4. Note that these lower bounds are existential, and not computational.

In Section 5, we consider the case where $\mathcal{M}$ is the set of nodes of an unweighted graph, and d measures the shortest distance between two nodes on the graph. When the graph is a tree, we show that an exact proportionally fair clustering necessarily exists, and provide an efficient algorithm to find one. When the graph is arbitrary, but $k \geq n/2$ clusters need to be placed, we show that a proportionally fair clustering again necessarily exists and can be computed efficiently. Whether an exact proportionally fair clustering exists for all graphs remains an interesting open question.

Next, in Section 6, we show that for $d = L^2$ and $\mathcal{M} = \mathbb{R}^t$ for $t \geq 2$, checking whether a proportional clustering exists is NP-hard. When t is large, even implementing the greedy capture algorithm is NP-hard. However, this problem becomes efficiently solvable when t is constant, and when t is large, using a PTAS for an important sub-routine of greedy capture, we can efficiently compute a $2 \cdot (1 + \epsilon)$-proportionally fair clustering for any fixed $\epsilon > 0$.

Finally, in Section 7, we consider the problem of generalization: would a clustering that is proportionally fair with respect to samples drawn from $\mathcal{N}$ remain (approximately) proportionally fair with respect to the entire set $\mathcal{N}$? Chen et al. provide a positive answer for the case when $\mathcal{M}$ is finite. Using the framework of VC dimension, we show that, when $d = L^2$, the answer remains positive even when $\mathcal{M} = \mathbb{R}^t$.

2 Preliminaries

Let $\mathcal{N}$ be a set of n data points (or agents), which lie in a metric space $(\mathcal{X}, d)$, where $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a distance function satisfying the triangle inequality. For most of this work, we consider the case where $\mathcal{X} = \mathbb{R}^t$ for some $t \in \mathbb{N}$, but in Section 5, we consider the case where $\mathcal{X}$ is the set of nodes of a graph. We also focus on special distance functions such as the Euclidean distance (L^2), the Manhattan distance (L^1), and the L^∞ distance. Let $\mathcal{M} \subseteq \mathcal{X}$ be a the set of locations where cluster centers can be placed. In this work, we focus on the case where $\mathcal{M} = \mathcal{X}$. Let $k \in \mathbb{N}$. We denote by $[\mathcal{M}]^k$ the set of all subsets of M of size k. A k-clustering is a set $X \in [\mathcal{M}]^k$ containing k locations for cluster centers. We refer to each $x \in X$ as an *open* cluster center.

The cost to agent $i \in \mathcal{N}$ induced by a cluster center $x \in \mathcal{M}$ is the distance $d(i, x)$, and the cost to agent $i \in \mathcal{N}$ induced by a k-clustering X is the minimum distance from i to any cluster center, i.e., $d(i, X) \triangleq \min_{x \in X} d(i, x)$. Agent i is interested in minimizing her cost.

A set of points $S \subseteq \mathcal{N}$ containing at least $\lceil n/k \rceil$ is entitled at least one cluster center. If they can find a new cluster center that is better for each of them, we consider it a violation of fairness. Formally:

▶ **Definition 1.** *Given $\rho \geq 1$, we say a k-clustering $X \in [\mathcal{M}]^k$ is ρ-proportionally fair if there is no $S \subseteq \mathcal{N}$ with $|S| \geq \lceil n/k \rceil$ and $y \in \mathcal{M}$ such that $\rho \cdot d(i, y) < d(i, X)$ for all $i \in S$. If such a coalition S and point y exist, we refer to S as a* blocking coalition *and y as the center that they* deviate *to. When $\rho = 1$, we simply call this proportional fairness.*

The reason to define ρ-proportionally fair solutions for $\rho > 1$ is because (exactly) proportionally fair solutions may not exist [8].

In Section 5, we consider the problem of proportional fairness on graphs. Specifically, given an undirected graph $G = (V, E)$, we assume that $\mathcal{N} \subseteq V$, $\mathcal{M} = V$ (i.e. every node in the graph is a feasible cluster center), and the distance between two nodes $u, v \in V$, denoted by $d(u, v)$, is the length of the shortest path connecting them. Notice that d satisfies the triangle inequality.

3   Greedy Capture

In this section, we study the greedy capture algorithm defined by Chen et al. [8]. Put succinctly, the algorithm starts with $X = \emptyset$. It grows a ball at every location in $\mathcal{M}$ at the same rate. As soon as a ball contains at least $\lceil n/k \rceil$ points, the corresponding center is added to X and all the points covered by the ball are removed. As balls continue growing, balls centered at previously added locations in X also continue growing with them and any new points covered by such balls are immediately removed. We refer the reader to the work of Chen et al. for full description of the algorithm. They show that for any metric space, greedy capture is guaranteed to find a $(1 + \sqrt{2})$-proportionally fair clustering.

We begin by providing a refined analysis of greedy capture by expressing the approximation ratio in terms of a characteristic of the metric we call the Apollonius radius.

▶ **Definition 2.** *Given $\rho \geq 1$, the ρ-Apollonius radius of a metric $(\mathcal{X}, d)$ is defined as $A_{\mathcal{X},d}(\rho) = \sup_{x,y \in \mathcal{X}} \Delta(\rho, x, y)/d(x, y)$, where $\Delta(\rho, x, y)$ is the radius of the smallest ball centered at some point in $\mathcal{X}$ that contains the entire set $\{p \in \mathcal{X} : \rho \cdot d(p, y) \leq d(p, x)\}$.*

The reason that we term it the Apollonius radius is because the renowned Greek geometer Apollonius of Perga was famously interested in the set $\{p \in \mathbb{R}^t : \rho \cdot d(p, y) \leq d(p, x)\}$ for $d = L^2$, and showed that this set is a ball already when $\rho > 1$. This special structure of L^2 is what will allow us to achieve a better approximation guarantee for it.

▶ **Theorem 3.** *For any metric $(\mathcal{X}, d)$ and $\mathcal{M} = \mathcal{X}$, greedy capture finds a ρ-proportionally fair clustering, where $\rho \geq 1$ is the smallest positive number satisfying $A_{\mathcal{X},d}(\rho) \cdot \frac{\rho+1}{\rho} \leq 1$.*

Proof. Let X be the clustering returned by the algorithm. Suppose that X is not ρ-proportionally fair for some ρ. Then, there exist $S \subseteq \mathcal{N}$ with $|S| \geq \lceil n/k \rceil$ and $y \in \mathcal{M}$ such that $\rho \cdot d(j, y) < d(j, X)$ for all $j \in S$.

Note that this implies $y \notin X$. Let i be the first point in S that was covered during the execution of greedy capture; suppose it was covered by a ball located at $x \in X$.

Note that for each $j \in S$, we have $\rho \cdot d(j, y) < d(j, X) \leq d(j, x)$. Hence, $S \subseteq \{p \in \mathcal{X} : \rho \cdot d(p, y) < d(p, x)\}$. Hence, by definition of ρ-Apollonius radius, there exists a ball of radius at most $A_{\mathcal{X},d}(\rho) \cdot d(x, y)$ that contains all points in S. Since i was the first point in S covered by the ball centered at x, we must have $d(i, x) \leq A_{\mathcal{X},d}(\rho) \cdot d(x, y)$, otherwise greedy capture would have selected the ball of radius at most $A_{\mathcal{X},d}(\rho) \cdot d(x, y)$ covering all points of S prior to i being covered by the ball centered at x. Further, since $i \in S$, we also have $d(i, y) < d(i, x)/\rho$. Using the triangle inequality, we get

$$d(x, y) \leq d(i, x) + d(i, y) < d(i, x) \cdot \frac{\rho + 1}{\rho} \leq A_{\mathcal{X},d}(\rho) \cdot \frac{\rho + 1}{\rho} \cdot d(x, y) \Rightarrow A_{\mathcal{X},d}(\rho) \cdot \frac{\rho + 1}{\rho} > 1.$$

We have proved that if X is not ρ-proportionally fair, then $A_{\mathcal{X},d}(\rho) \cdot \frac{\rho+1}{\rho} > 1$. Hence, whenever $A_{\mathcal{X},d}(\rho) \cdot \frac{\rho+1}{\rho} \leq 1$, we have that X is ρ-proportionally fair. ◄

Let us argue that Theorem 3 is in fact a refinement of the $(1 + \sqrt{2})$-approximation proved by Chen et al. [8] that holds for all metrics.

▶ **Theorem 4.** *For any metric $(\mathcal{X}, d)$, the ρ-Apollonius radius is $A_{\mathcal{X},d}(\rho) \leq \frac{1}{\rho-1}$. Hence, greedy capture finds a $(1 + \sqrt{2})$-proportionally fair clustering for every metric.*

Proof. Consider the set $\{p \in \mathcal{X} : \rho \cdot d(p, y) \leq d(p, x)\}$. For any point p in this set, we have that $\rho \cdot d(p, y) \leq d(p, x) \leq d(p, y) + d(x, y)$ by the triangle inequality. Hence, $d(p, y) \leq d(x, y) \cdot \frac{1}{\rho-1}$ for all p in the set. Thus, a ball centered at $y \in \mathcal{M}$ with radius $\frac{d(x,y)}{\rho-1}$ certainly covers the entire set. Hence, $A_{\mathcal{X},d}(\rho) \leq \frac{1}{\rho-1}$. Next, for $\rho = 1 + \sqrt{2}$, we have

$$A_{\mathcal{X},d}(\rho) \cdot \frac{\rho + 1}{\rho} \leq \frac{\rho + 1}{\rho \cdot (\rho - 1)} = 1.$$

Hence, by Theorem 3, greedy capture finds a $(1 + \sqrt{2})$-proportionally fair clustering. ◄

Next, we show that for $d = L^2$, the ρ-Apollonius radius is slightly better, leading to a better 2-approximation guarantee for greedy capture.

▶ **Theorem 5.** *For the metric space $(\mathbb{R}^t, L^2)$, where $t \in \mathbb{N}$, the ρ-Apollonius radius is $A_{\mathbb{R}^t, L^2}(\rho) \leq \frac{\rho}{\rho^2-1}$, and hence, greedy capture finds a 2-proportionally fair clustering.*

Proof. For the L^2 norm in a Euclidean space, it is well-known that given $x, y \in \mathbb{R}^t$ and $\rho > 1$, the set of points $\{p \in \mathbb{R}^t : \rho \cdot d(p, y) \leq d(p, x)\}$ is a ball of radius $d(x, y) \cdot \frac{\rho}{\rho^2-1}$. This is a simple algebraic exercise; its two-dimensional variant was known to Apollonius himself, after whom the result is named (the derivation is widely available online, e.g., see [11]). This immediately implies that $A_{\mathbb{R}^t, L^2}(\rho) \leq \frac{\rho}{\rho^2-1}$.

Now, we have that

$$A_{\mathcal{X},d}(\rho) \cdot \frac{\rho + 1}{\rho} \leq \frac{\rho}{\rho^2 - 1} \cdot \frac{\rho + 1}{\rho} = \frac{1}{\rho - 1}.$$

This quantity is at most 1 when ρ is at least 2. Hence, by Theorem 3, greedy capture finds a 2-proportionally fair clustering for this metric. ◄

The obvious next question then is whether this refinement also provides an improved approximation bound for other distance metrics. Unfortunately, for two other prominent distance metrics, L^1 and L^∞, the answer is no. We show this by providing a direct counterexample where greedy capture finds a clustering that is no better than $(1 + \sqrt{2})$-proportionally fair. The proof of the next result appears in the full version of the paper.

▶ **Theorem 6.** *For the metric space $(\mathbb{R}^t, d)$ where $t \geq 2$ and $d \in \{L^1, L^\infty\}$, and $\mathcal{M} = \mathbb{R}^t$, there exists an example in which the clustering produced by greedy capture is not ρ-proportionally fair for $\rho < 1 + \sqrt{2}$.*

4   Universal Lower Bounds

In this section, we show lower bounds on approximation to proportional fairness that apply to all algorithms, as opposed to the lower bounds in the previous section that apply only to greedy capture. Chen et al. [8] show that when $\mathcal{N}$, $\mathcal{M}$, and the metric are arbitrary, ρ-proportional fairness cannot be guaranteed for $\rho < 2$. They also consider the special case where $\mathcal{N} = \mathcal{M}$, and prove a slightly weaker lower bound of 1.5. One question that they do not address is whether greedy capture provides better than $(1 + \sqrt{2})$-approximation in this special case; in the full version, we show that this is *not* the case.

In this section, we turn our attention to the case of our interest: $\mathcal{M} = \mathcal{X} = \mathbb{R}^t$ and $d \in \{L^1, L^2, L^{\inf}\}$. When $t = 1$, it is easy to notice that an exactly proportionally fair clustering always exists.[1] When $t \geq 2$, we provide a lower bound of $2/\sqrt{3}$ for $d = L^2$ and a lower bound of 1.4 for $d \in \{L^1, L^\infty\}$.

▶ **Theorem 7.** *For the metric space $(\mathbb{R}^t, L^2)$ where $t \geq 2$ and $\mathcal{M} = \mathbb{R}^t$, there is an example in which no clustering is ρ-proportionally fair for $\rho < 2/\sqrt{3} \approx 1.155$.*

Proof. Once again, we set $t = 2$ without loss of generality. Consider an instance in which $|\mathcal{N}| = 6$ and $k = 3$. Suppose $|\mathcal{N}|$ consists of two isomorphic sets of 3 points each, where each set of 3 points forms an equilateral triangle of length 1 and the two sets are sufficiently far from each other. Then, by the pigeonhole principle, under any clustering X, at least one set of 3 points, say $\{p_1, p_2, p_3\}$, must derive their costs from a single cluster center x.

Let a denote the circumcenter of their triangle. Then, $d(a, p_1) = d(a, p_2) = d(a, p_3) = 1/\sqrt{3}$. Hence, $d(a, p_1) + d(a, p_2) + d(a, p_3) = \sqrt{3}$. Notice that in an equilateral triangle, the circumcenter is also the Fermat point, which minimizes the sum of distances from the three vertices. Hence, for the cluster center x, we have $d(x, p_1) + d(x, p_2) + d(x, p_3) \geq \sqrt{3}$. Without loss of generality, assume $d(x, p_1) \geq d(x, p_2) \geq d(x, p_3)$. Then, $d(x, p_1) + d(x, p_2) \geq 2/\sqrt{3}$.

Now, p_1 and p_2 can deviate, choose a location y on the line joining p_1 and p_2 such that $d(y, p_1)/d(y, p_2) = d(x, p_1)/d(x, p_2)$. Since $d(y, p_1) + d(y, p_2) = d(p_1, p_2) = 1$, this reduces the cost to each point by a factor of $2/\sqrt{3}$. Hence, the clustering is not ρ-proportionally fair for $\rho < 2/\sqrt{3}$. ◀

Note that the lower bound of 1.155 is significantly lower than the upper bound of 2 obtained by greedy capture for L^2 as shown in Theorem 5. Closing the gap is an interesting open question. Next, we show a lower bound for L^1 and L^∞. The proof appears in the full version of the paper.

▶ **Theorem 8.** *For the metric space $(\mathbb{R}^t, d)$, where $t \geq 2$ and $d \in \{L^1, L^\infty\}$, and $\mathcal{M} = \mathbb{R}^t$, there is an example in which no clustering is ρ-proportionally fair for $\rho < 1.4$.*

5   Clustering in Graphs

In this section, we consider the special case where the metric space $(\mathcal{X}, d)$ is induced by an undirected graph $G = (V, E)$. Specifically, we let $\mathcal{X} = V$ be the set of nodes of the graph, and assume that $d(x, y)$ measures the length of the shortest path between nodes x and y. As in the previous sections, we restrict our attention to the $\mathcal{M} = \mathcal{X}$ case, i.e., when every point of the metric space is a feasible cluster center.

---

[1] For instance, opening a cluster at every n/k-th data point from left to right is proportionally fair.

Graphs are an important special case since clustering and facility location in graphs and networks are very well studied. However, while objectives such as truthfulness [2] and social welfare maximization (or social cost minimization) [13] have received significant attention, fairness has not.

We study proportional fairness for this setting. Our first result shows that when the graph $G = (V, E)$ is a tree, an exact proportionally fair clustering always exists, and can be computed by an efficient algorithm. Intuitively, the algorithm works as follows. We first root the tree at an arbitrary node r to obtain a rooted tree (G, r). We denote with h the height of the rooted tree, and with level(x) the height of node x relative to the root r (with level$(r) = 1$). Let ST(x) denote the subtree of node x (i.e. the set of nodes v such that with level$(v) \geq$ level(x) and the unique path from v to r contains x). The algorithm starts from the highest level (the leaves), opens a center every time it finds a node whose subtree contains at least $\lceil n/k \rceil$ nodes, and deletes this subtree from the graph. At the end, the cost to each node is still defined using the closest node at which a center is opened by the algorithm.

🟨 **Algorithm 1** Proportionally Fair Clustering for Trees.

1: Root the tree G at an arbitrary node r
2: $X \leftarrow \emptyset$
3: $G^h \leftarrow G$
4: **for** $\ell = h$ to 1 **do**
5:     $G^{\ell-1} \leftarrow G^\ell$
6:     **for** every $x \in V$ with level$(x) = \ell$ and $|\text{ST}(x)| \geq \lceil n/k \rceil$ **do**
7:         $X \leftarrow X \cup \{x\}$
8:         $G^{\ell-1} \leftarrow G^{\ell-1} \setminus \text{ST}(x)$
9: **if** $G^0 \neq \emptyset$ **then**
10:     $X \leftarrow X \cup \{r\}$
      return X

▶ **Theorem 9.** *Let $G = (V, E)$ be an undirected tree, (V, d) be the metric induced by G, $\mathcal{N} \subseteq V$, $\mathcal{M} = V$, and $k \in \mathbb{N}$. Then, Algorithm 1 yields a proportionally fair clustering.*

Proof. Let X be the clustering returned by Algorithm 1. First, we notice that X contains at most k centers. This is because every time the algorithm opens a center in the for loop, it deletes at least $\lceil n/k \rceil$ nodes from the graph. Hence, the for loop can add at most k centers. If it adds exactly k centers, then the remaining graph must be empty, and the if condition is not executed. If it adds at most $k-1$ centers, then even if the root node is added later, there would be at most k centers.

Next, suppose for contradiction that X is not proportionally fair. Hence, there exists a set $S \subseteq V$ with $|S| \geq \lceil n/k \rceil$ and $y \in V$ such that $d(i, y) < d(i, X)$ for all $i \in S$.

For each node $i \in V$, define $p(i)$ to be its closest ancestor in X (i.e. $p(i) \in X$ and $i \in \text{ST}(p(i))$, and $p(i)$ is the node of maximum level satisfying these two conditions). Note that because the algorithm always opens a center at the root node, $p(i)$ is well-defined for each node i.

Further, note that for each $i \in S$, $d(i, X) \leq d(i, p(i))$. And for nodes $j \notin \text{ST}(p(i))$, $d(i, j) > d(i, p(i))$. Hence, the cost to i can only reduce if the deviating center is in $\text{ST}(p(i))$. We now consider two cases.

- **Case 1:** $\exists i, i' \in S : p(i) \neq p(i')$. First, suppose that $p(i)$ and $p(i')$ are siblings (i.e. $\text{ST}(p(i)) \cap \text{ST}(p(i')) = \emptyset$). As i can only improve if the deviating center is in $\text{ST}(p(i))$ and i' can only improve if the deviating center is in $\text{ST}(p(i'))$, we obtain that no deviating center y can reduce the cost to i and i' simultaneously, which is a contradiction.

Next, suppose that $p(i) \in \mathrm{ST}(p(i'))$. Then, we must have $y \in \mathrm{ST}(p(i))$, otherwise the cost to i would not reduce. But then, $d(i', p(i)) \leq d(i', y)$. Hence, the cost to i' does not reduce due to y, which is also a contradiction.

The remaining case of $p(i') \in \mathrm{ST}(p(i))$ is symmetric to the last case.

■ **Case 2:** $\forall i \in S, p(i) = p^*$. Let $O = X \cap \mathrm{ST}(p^*) \setminus \{p^*\}$ be the set of open centers in $\mathrm{ST}(p^*)$ except p^* itself. Note that by definition of p^*, we have that if $i \in S$, then $i \notin \mathrm{ST}(o)$ for any $o \in O$.

This implies that if $y \in \mathrm{ST}(o)$ for some $o \in O$, then for every point $i \in S$, we have $d(i, y) \geq d(i, o) \geq d(i, X)$, meaning that y would not reduce the cost to any point in S. Hence, $y \in \mathrm{ST}(p^*) \setminus \cup_{o \in O} \mathrm{ST}(o)$.
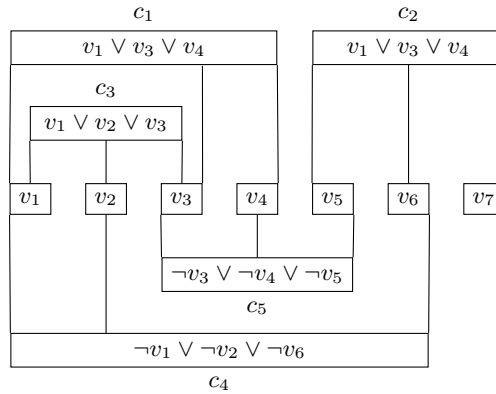
In other words, if center p^* was opened in the iteration with index ℓ (let $\ell = 0$ if p^* is the root node that was opened outside of the for loop), then $S \cup \{y\} \in G^\ell$ (any point from S or y could not have been deleted in any previous iteration). However, for y to reduce the cost to each $i \in S$, we must have $S \subseteq \mathrm{ST}(y)$. However, then, y is a node of higher level than p^* that still contains at least $|S| \geq \lceil n/k \rceil$ points, so it must have been removed in a previous iteration. This is the desired contradiction.

This concludes the proof.   ◀
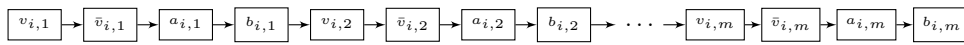
We remark that Theorem 9 identifies the broadest class of interesting metrics to date for which an exact proportionally fair clustering is known to always exist. This raises an immediate question: what about graphs that are not trees? We can consider the universal lower bound for the $(\mathbb{R}^2, L^1)$ metric from Theorem 8. If we construct a very dense grid graph (in which the shortest path distance mimics the L^1 distance in the plane) in the relevant region of $\mathbb{R}^2$ from that example, we can derive the same lower bound of 1.4 on the approximation ratio to proportional fairness for graphs. Whether better lower bounds exist is an open question. Similarly, the $1 + \sqrt{2}$ upper bound of greedy capture holds for all metrics, including those induced by a graph. However, whether its approximation ratio or running time can be improved for graphs, possibly with special properties such as planarity, remains to be seen.

It is worthwhile noting the special case of $\mathcal{N} = \mathcal{M} = V$, i.e., when every node of the graph is also a data point (besides being a feasible cluster center location). In this case, we do not know whether an exact proportionally fair clustering always exists, and leave this as an interesting open question. That said, we do note that if G is connected and we want to place a large number of clusters $k \geq n/2$, then it can be shown that a proportionally fair clustering exists. This is because a dominating set[2] of any size $k \geq n/2$ is guaranteed to exist in a graph with n nodes [26] and can be computed efficiently [18]. If nodes in such a set are chosen as the cluster centers, then every node in the graph already has cost at most 1. So to deviate, all nodes in the blocking coalition must achieve cost 0. However, since the blocking coalition must contain at least $\lceil n/k \rceil \geq 2$ nodes, this is impossible. Thus, the questions of existence and computation of a proportionally fair clustering in general graphs with $\mathcal{N} = \mathcal{M} = V$ become trivial when $k \geq n/2$, but remain open when $k < n/2$.

---

[2] A set of nodes is called a dominating set if every node in the graph is either in this set or adjacent to a node in this set.

Figure 1 An example of labelling clauses and variables.

Figure 2 The variable gadget corresponding to variable v_i. Directed edges indicate the closest neighbour of each node.

6 Computational Aspects

In this section, we consider computational aspects of two problems: the problem of checking whether a proportionally fair clustering exists, and the problem of implementing the greedy capture algorithm when $\mathcal{M} = \mathbb{R}^t$. We begin by considering the former problem. The full proof appears in the full version of the paper, but we provide a sketch here.

▶ **Theorem 10.** *Given $\mathcal{X} = \mathbb{R}^2$, finite $N \subset \mathcal{X}$, finite $M \subset \mathcal{X}$, $k \in \mathbb{N}$, and $d = L^2$, checking whether a proportionally fair clustering exists is NP-hard.*

Proof Sketch. We first show a reduction which creates an instance of proportionally fair clustering with $n/k = 2$, and later show how to extend this to the case where n/k is any even integer. We use a polynomial-time reduction from the planar monotone rectilinear 3-SAT problem, where each clause c_j consists of only positive or only negative literals, each variable v_i is represented by a rectangle on the x-axis and each positive (resp. negative) clause is represented by a rectangle above (resp. below) the x-axis with three vertical lines or legs to its three variables. The graph that connects clauses to literals is planar. Figure 1 shows what a planar monotone rectilinear 3-SAT instance looks like.

Let I be an instance of a planar monotone rectilinear 3-SAT which consists of l boolean variables and m monotone clauses. Given I, we construct an instance I' of proportionally fair clustering with $|\mathcal{N}| = \Theta(lm^2)$ and $\mathcal{M} = \Theta(lm^2)$ such that I is satisfiable if and only if there exists a proportionally fair clustering in I'.

First, for each variable v_i, we construct a variable gadget which contains points $v_{i,j}$, $\bar{v}_{i,j}$, $a_{i,j}$ and $b_{i,j}$ for $j \in [m]$. They are all on the line, and belong to both $\mathcal{N}$ and $\mathcal{M}$. This gadget is shown in Figure 2. The point $v_{i,j}$ (resp. $\bar{v}_{i,j}$) corresponds to the positive (resp. negative) literal of v_i, specifically reserved for clause c_j (whether or not it appears in that clause). We set the distances in a way that the closest node to any node is the node on its right, while the closest node to the last node $b_{i,m}$ is its previous node $a_{i,m}$.

All variable gadgets are located on the x-axis in such a way that the gadget of variable v_i is on the left of the gadget of variable v_{i+1}, and from left to right, the distances between two adjacent variable gadgets are slightly decreasing. (The exact construction is provided in the full proof.)

Next, we construct two clause gadgets, a basic and an auxiliary one, for each clause. The auxiliary gadget of each clause is sufficiently far from all other gadgets in the construction. It consists of 3 points and 3 feasible centers such that at least two centers must be placed within each such gadget, otherwise, ρ-proportional fairness is violated for all $\rho < 1.214$. The purpose of this gadget is to make fewer than a proportional number of centers available for the rest of the construction.

In the basic gadget of a clause, we add points that form a rectangle with three legs. The interaction of this gadget with variable gadgets is shown in Figure 3. Let y_j be the y-coordinate where we place all the points that consist the virtual rectangle of c_j. We choose the values of $y_i, i \in [m]$, such that each point in a basic gadget has as closest neighbour a point in the same gadget. Let c_j be a positive clause which contains the positive literals of variables v_q, v_s and v_t with $q < s < t$ (respectively, if the clause is negative). We add a virtual rectangle in the interval $[(v_{q,j}, y_j), (v_{t,j}, y_j)]$ and add the three virtual legs that are vertical at x-coordinates $v_{q,j}$, $v_{s,j}$, and $v_{t,j}$, respectively. First, we place $2m(t-s)$ points in the interval $[(\bar{v}_{s,j}, 0), (v_{t,j}, 0)]$. Denote these points as $r_{j,k}^1$, $k \in [1, 2m(t-s)]$, where the x-coordinate of $r_{j,k}^1$ is less than the x-coordinate of $r_{j,k+1}^1$, and we set the distances such that $r_{j,k}^1$ has as its closest neighbor $r_{j,k+1}^1$. Second, we place in a similar way $2m(s-q)$ points in the interval $[(v_{q,j}, y_j), (a_{s,j-1}, y_j)]$ (or $[(v_{qj}, y_j), (a_{s-1,m}, y_j)]$ if $j = 1$). We denote these points as $r_{j,k}^2$, $k \in [1, 2m(s-q)]$, where the x-coordinate of point $r_{j,k}^2$ is larger than the x-coordinate of point $r_{j,k+1}^2$, and we set the distances such that $r_{j,k}^2$ has as its closest neighbor $r_{j,k+1}^2$.
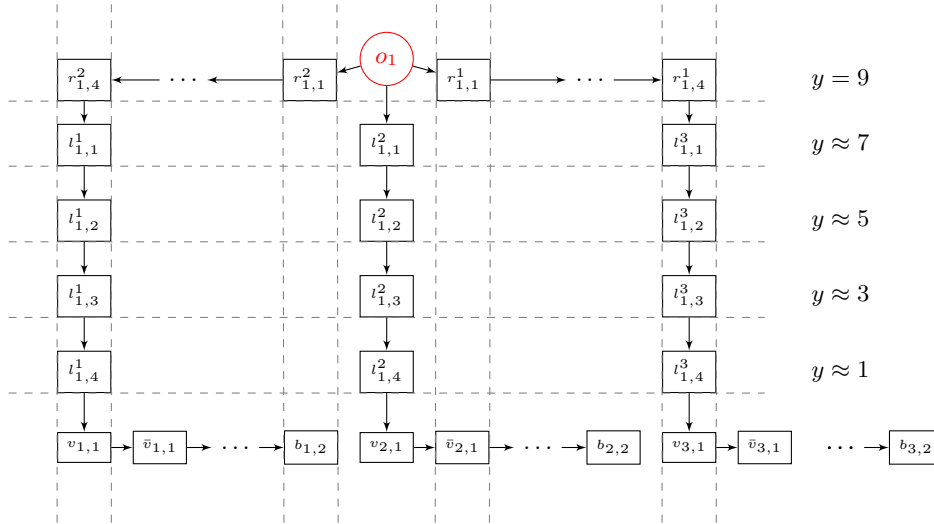
It remains to construct the legs of each clause. First, we place $2n_j$ points (see the full version for the exact value of n_j), denoted by $l_{j,k}^1$, $k \in [1, 2n_j]$, with x-coordinate equal to $v_{q,j}$ and y-coordinate less than y_j. Specifically, we locate $l_{j,1}^1$ in a position such that the leftmost point in the rectangle of c_j has as its closest neighbor $l_{j,1}^1$, the closest neighbor of every $l_{j,k}^1$, $k \in [1, 2n_j - 1]$, is $l_{j,k+1}^1$, and the closest neighbor of $l_{j,2n_j}^1$ is $v_{q,j}$. For the remaining legs, we add points at exactly the same y-coordinates of the points in the first leg, but with x-coordinates equal to $v_{s,j}$ and $v_{t,j}$. Denote the points of the middle and the right leg as $l_{j,k}^2$ and $l_{j,k}^3$, $k \in [1, 2tn_j]$, respectively.

Lastly, for each clause we add one more point, denoted by o_j in a location such that it is the circumcenter of the triangle with nodes $r_{j,1}^1$, $r_{j,1}^2$, and $l_{j,1}^2$, and these are the (tied) closest neighbors of o_j. Figure 3 shows this entire construction for an example instance which consists of 2 positive clauses (only one of which is shown in the figure) and 3 variables.

In this construction, note that each clause gadget (the union of basic and auxiliary gadgets) has an even number of points equal to $2n_j' + 4$ (for some n_j'; see the full version for details), and each variable gadget has $4m$ of points. Hence, we choose $k = 2ml + \sum_{j=1}^m (n_j' + 2m)$, so that $n/k = 2$. Now, we are ready to prove that I is satisfiable if and only if there exists a proportionally fair clustering in I'.

Note that in each variable gadget we need at least $2m$ cluster centers in order for the clustering to be proportionally fair. This is because every pair of adjacent points can deviate if neither of them is a cluster center. There are only two ways to place exactly $2m$ centers: we can either open centers at $v_{i,j}$ and $a_{i,j}$ for all $j \in [4m]$, or open centers at $\bar{v}_{i,j}$ and $b_{i,j}$ for $j \in [4m]$. The first choice corresponds to an assignment where x_i is set to true, while the second corresponds to an assignment where x_i is set to false.

As we mentioned earlier, the auxiliary gadget of each clause needs at least two cluster centers placed within it, otherwise a proportionally fair clustering cannot exist. This leaves n_j' centers for every basic clause gadget. However, each basic gadget needs at least n_j' centers. To see this, notice that from point $r_{j,1}^2$ to the last point of the left leg, we need to add one center at every other point. Similarly, from point $r_{j,1}^1$ to the last point of the right leg, we

Figure 3 The rectangle and the legs of clause $c_1 = v_1 \vee v_2 \vee v_3$ in an instance with $m = 2$ and $l = 3$. A directed edge from one node points to the closest neighbor of the node.

also need a center at every other point. The same holds for the middle leg. As these are $2n'_j$ points in total, this requires at least n'_j centers in an alternating pattern. Notice that at least one of $r^1_{j,1}$, $r^2_{j,1}$ and $l^2_{j,1}$ must also be an open center, otherwise o_j can deviate with one of them. This is possible only if at least one of the corresponding variable nodes is an open center. The reason is that the last point of at least one leg is not an open center and this point should not want to deviate with its closest node (which corresponds to the literal of the clause). This happens if and only if this node is a center, and so the clause is satisfied.

This shows that I is satisfiable if and only if I' admits a proportionally fair clustering.   ◄

Next, we consider implementing the greedy capture algorithm when $\mathcal{M} = \mathbb{R}^t$. Note that because the naive description of greedy capture requires simultaneously growing a ball from each point in $\mathcal{M}$, it is easy to implement when $\mathcal{M}$ is finite (as shown by Chen et al. [8]) but difficult when $\mathcal{M}$ is infinite.

To avoid this issue, one may consider restricting the set of feasible cluster center locations to a finite set $\mathcal{M}'$, and hope that running greedy capture (or any algorithm for that matter) on this finite set still yields a clustering that is approximately fair with respect to the original infinite $\mathcal{M}$. To that end, we show that restricting to $\mathcal{M}' = \mathcal{N}$ (i.e. choosing cluster centers from the set of data points) can only worsen the approximation factor of an algorithm by a factor of at most 2. The proof appears in the full version.

▶ **Theorem 11.** *For any metric space $(\mathcal{X}, d)$, $k \in \mathbb{N}$, $\mathcal{N}, \mathcal{M} \subseteq \mathcal{X}$, and $\rho \geq 1$, any ρ-proportionally fair k-clustering with respect to $(\mathcal{N}, \mathcal{M}' = \mathcal{N})$ is 2ρ-proportionally fair with respect to $(\mathcal{N}, \mathcal{M})$.*

For the metric $(\mathbb{R}^t, L^2)$, where $t \in \mathbb{N}$, recall that greedy capture produces a 2-proportionally fair clustering when $\mathcal{M} = \mathbb{R}^t$. Hence, Theorem 11 implies that running it on $\mathcal{M}' = \mathcal{N}$, for finite $\mathcal{N}$, would yield a 4-proportionally fair clustering in polynomial time. However, for this special case, we can in fact efficiently achieve $2 + \epsilon$ approximation for any constant $\epsilon > 0$, as we discuss below.

Let us consider the first cluster center added by greedy capture when $\mathcal{M} = \mathbb{R}^t$. It is the center of a smallest ball that contains at least n/k of the n given points. The problem of finding the smallest ball containing at least p of n given points is a well-studied problem in computational geometry. This is known to be NP-hard [31], which is what makes implementing greedy capture hard when $\mathcal{M} = \mathbb{R}^t$.

However, when t is constant, the problem is known to be solvable efficiently [31]; we show that in this case, greedy capture can also be implemented efficiently, and we can achieve the 2-approximation from Theorem 5.

Further, even when t is large, the problem admits a PTAS [31]. We show that this PTAS can be used to approximately implement greedy capture, and the approximation ratio only slightly worsens. The proof of the next result appears in the full version.

▶ **Theorem 12.** *Let $t \in \mathbb{N}$, finite $\mathcal{N} \subset \mathbb{R}^t$, and $k \in \mathbb{N}$ be given as input. Suppose $\mathcal{M} = \mathbb{R}^t$ and $d = L^2$. Then, the following hold.*

1. *The clustering returned by greedy capture algorithm cannot be computed in polynomial time unless $P = NP$.*
2. *If t is constant, then it can be computed in polynomial time.*
3. *Even if t is not constant, for any constant $\epsilon > 0$, there exists a polynomial-time algorithm which finds a $(2 + \epsilon)$-proportionally fair clustering.*

7 Learning Fair Clustering

A key concern in machine learning is generalization. In our context, the question is whether a clustering that is (approximately) proportionally fair with respect to random samples taken from an underlying population would remain (approximately) proportionally fair with respect to the whole population. A positive answer to this question could be useful in two ways.

First, sometimes we may have access only to samples from an underlying population. In this case, we can rest assured that by computing a clustering that is fair with respect to the samples, it is also fair with respect to the population. Second, even if the entire population is known, it may be very large. As we noticed in Section 6, finding a proportionally fair clustering or even running the greedy capture algorithm is NP-hard; thus, these tasks may be infeasible for a large population. However, it may be possible to do so on a smaller sample taken from the population, which is where the generalization guarantee can be useful.

Chen et al. [8] show that generalization indeed holds for proportional fairness. Specifically, they define the following relaxation of ρ-proportional fairness.

▶ **Definition 13.** *We say that a k-clustering X is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$ if for all $S \subseteq \mathcal{N}$ with $|S| \geq |\mathcal{N}| \cdot (1 + \epsilon)/k$ and all $y \in \mathcal{M}$, there exists at least one $i \in S$ such that $\rho \cdot d(i, y) \geq d(i, X)$.*

Chen et al. show that if $N \subseteq \mathcal{N}$ is a uniformly random sample of size $|N| = \Omega\left(\frac{k^3}{\epsilon} \ln \frac{|\mathcal{M}|}{\delta}\right)$, and if X is ρ-proportionally fair with respect to N, then X is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$ with probability at least $1 - \delta$.

Unfortunately, this bound depends on $|\mathcal{M}|$, and breaks down when $|\mathcal{M}|$ is infinite, which is the focus of our work. We establish a stronger guarantee that does not depend on $|\mathcal{M}|$ by utilizing the framework of VC dimension [33] for binary classifiers. First, we show that there is a natural family of binary classifiers associated with a given clustering.

▶ **Definition 14.** *Given a set of points $\mathcal{N}$, a k-clustering $X \in [\mathcal{M}]^k$, and $y \in \mathcal{M}$, define the binary classifier $h_{X,y} : \mathcal{N} \to \{0, 1\}$ such that $h_{X,y}(i) = 1$ if and only if $\rho \cdot d(i, y) < d(i, X)$. Define the "error" of this classifier on a set of points $S \subseteq \mathcal{N}$ as $err_S(h_{X,y}) = (1/|S|) \cdot \sum_{i \in S} h_{X,y}(i)$.*

Intuitively, $h_{X,y}(i) = 1$ if and only if i can be part of a coalition that complains about the unfairness of X by demonstrating y as a location that provides them ρ-improvement. The use of the term "error" may be confusing. Unlike in traditional classification context, where there is a true classifier and the error is measured in terms of the fraction of points on which a given classifier differs from the true classifier, in our case the "error" is simply the fraction of points that can deviate. One can equivalently think of the "true classifier" as the one that outputs 0 on every point.

Note that X is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$ if and only if $\mathrm{err}_{\mathcal{N}}(X, y) \leq \frac{1+\epsilon}{k}$ for all $y \in \mathcal{M}$. Our goal is to show that given a sufficiently large random sample $N \subseteq \mathcal{N}$, if we have a clustering X that is ρ-proportionally fair with respect to N, then it is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$ with high probability. However, note that we have no control over what X or y are. This is where the stronger "uniform convergence" guarantee – this establishes that $|\mathrm{err}_N(h_{X,y}) - \mathrm{err}_{\mathcal{N}}(h_{X,y})|$ is bounded for all X, y – becomes useful. Let us begin by introducing the VC dimension and a well-known uniform convergence guarantee that depends on the VC dimension.

▶ **Definition 15** (VC Dimension). *Let $\mathcal{N}$ be a set of points. Let $\mathcal{H}$ be a family of binary classifiers over $\mathcal{N}$. We say that $\mathcal{H}$ shatters $S \subseteq \mathcal{N}$ if for every labeling $\ell : S \to \{0, 1\}$, there exists a classifier $h \in \mathcal{H}$ such that $h(i) = \ell(i)$ for all $i \in S$. The VC dimension of $\mathcal{H}$, denoted $\dim^{\mathrm{VC}}(\mathcal{H})$, is the size of the largest S that can be shattered by $\mathcal{H}$.*

▶ **Proposition 16** ([29]). *Let $\mathcal{H}$ be a family of binary classifiers over a set of points $\mathcal{N}$. If $N \subseteq \mathcal{N}$ is a uniformly random sample with $|N| \geq \Omega\left(\frac{\dim^{\mathrm{VC}}(\mathcal{H}) + \ln(1/\delta)}{\epsilon^2}\right)$, then with probability at least $1 - \delta$, $|\mathrm{err}_N(h) - \mathrm{err}_{\mathcal{N}}(h)| \leq \epsilon$ for all $h \in \mathcal{H}$.*

We show that the family of classifiers $\{h_{X,y} | X \in [\mathcal{M}]^k, y \in \mathcal{M}\}$ has finite VC dimension when $\mathcal{M} = \mathbb{R}^t$ with finite t. This, along with Proposition 16, gives us the desired result.

▶ **Theorem 17.** *Fix $\epsilon, \delta > 0$, $\rho \geq 1$, $k, t \in \mathbb{N}$, and metric $(\mathbb{R}^t, d)$ where $d = L^2$. Let $\mathcal{N}$ be a set of points and $\mathcal{M} = \mathbb{R}^t$ be the set of feasible cluster centers. Let $N \subseteq \mathcal{N}$ be sampled uniformly at random with $|N| \geq \Omega\left(\frac{k^2 \cdot (tk \ln k + \ln(1/\delta))}{\epsilon^2}\right)$. Then, with probability at least $1 - \delta$, every k-clustering $X \in [\mathcal{M}]^k$ that is ρ-proportionally fair with respect to N is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$.*

Proof. Given a pair of points x and y, note that the set of points i such that $\rho d(i, y) \geq d(i, x)$ is a half-space in $\mathbb{R}^t$ when $\rho = 1$ and a ball in $\mathbb{R}^t$ when $\rho > 1$. Hence, given $X \in [\mathcal{M}]^k$, the set of points i satisfying $\rho d(i, y) \geq d(i, X)$ is the union of k half-spaces or balls in $\mathbb{R}^t$. It is known that the VC dimension of unions of k half-spaces or balls in $\mathbb{R}^t$ is $O(tk \ln k)$ [6].

Substituting this bound in Proposition 16, we get that $|N| \geq \Omega\left(\frac{k^2 \cdot (tk \ln k + \ln(1/\delta))}{\epsilon^2}\right)$ is sufficient to ensure that with probability at least $1 - \delta$, $\mathrm{err}_{\mathcal{N}}(h_{X,y}) \leq \mathrm{err}_N(h_{X,y}) + \epsilon/k$ for all X, y. In particular, when X is ρ-proportionally fair with respect to N, this ensures that with probability at least $1 - \delta$, X is ρ-proportionally fair to $(1 + \epsilon)$-deviations with respect to $\mathcal{N}$. ◀

While we do not formally consider the case of infinite set of points ($|\mathcal{N}| = \infty$) in this work, one can define $\mathcal{N}$ as a distribution over infinitely many points, and ask whether a probability mass of at least $1/k$ has a beneficial deviation. Note that Theorem 17 applies to this case as well because it does not depend on $|\mathcal{N}|$. On the other hand, note that Theorem 17 applies only for L^2 distance metric; deriving generalization bounds for other distance metrics remains an interesting challenge for future work.

8   Discussion

In this work, we advanced the study of proportionally fair clustering in a metric space, and focused on the case where the set of possible cluster center $\mathcal{M}$ is the entire (usually infinite) metric space. Our work leaves a number of open questions.

The most immediate question is to bridge the gap between our lower and upper bounds on the approximation ratio to proportional fairness from Sections 3 and 4. In particular, we believe that for L^2, the lower bound of $2/\sqrt{3}$ from Theorem 7 may be achievable. This specific number is reminiscent of Jung's theorem, which states that for L^2 distance in $\mathbb{R}^2$ (which is where the lower bound stems from), any set of points with diameter at most 1 is contained in a ball of radius at most $1/\sqrt{3}$. This could be useful in closing the gap for L^2. Section 5 leaves open an important question which is whether a proportionally fair clustering exists for all graphs when $\mathcal{N} = \mathcal{M}$ is the set of all nodes of the graph. Our hardness results from Section 6 and learnability results from Section 7 only apply to the L^2 distance because they use results and techniques from the literature that are only available for L^2. Deriving similar results for other distance metrics would be very interesting. While Chen et al. [8] show that achieving bounded approximation to proportional fairness is incompatible with achieving bounded approximation to three classic objective functions (k-center, k-means, and k-median), it would be interesting to study the tradeoff between fairness and these objectives in special cases (such as graphs).

In this paper, our focus was on the case where the set of possible cluster centers $\mathcal{M}$ is the entire metric space. In the case where the metric is induced by a graph $G = (V, E)$ (Section 5), it is also interesting to consider the case where $\mathcal{M} \subset V$. Does a proportionally fair clustering always exist for trees when $\mathcal{M} \subset V$, or for general graphs when $\mathcal{N} = \mathcal{M} \subset V$? These questions remain open.

More broadly, we find proportional fairness to be a very elegant fairness solution concept for clustering. Adapting this idea of proportional fairness to other machine learning settings such as regression or classification can lead to many avenues for future work.

References

1   Alekh Agarwal, Miroslav Dudík, and Zhiwei Steven Wu. Fair regression: Quantitative definitions and reduction-based algorithms. In *Proceedings of the 36thInternational Conference on Machine Learning (ICML)*, volume 97, pages 120–129, 2019.

2   Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. Strategyproof approximation of the minimax on networks. *Mathematics of Operations Research*, 35(3):513–526, 2010.

3   Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias: There's software used across the country to predict future criminals. *And it's biased against blacks. ProPublica*, 23, 2016.

4   Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, and Tal Wagner. Scalable fair clustering. In *Proceedings of the 36thInternational Conference on Machine Learning (ICML)*, pages 405–413, 2019.

5   Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. Fair algorithms for clustering. In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 4955–4966, 2019.

6   Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the vapnik-chervonenkis dimension. *Journal of the Association for Computing Machinery*, 36(4):929–965, 1989.

7 Toon Calders, Faisal Kamiran, and Mykola Pechenizkiy. Building classifiers with independency constraints. In *Proceedings of the IEEE International Conference on Data Mining Workshops (ICDM)*, pages 13–18, 2009.

8 Xingyu Chen, Brandon Fain, Charles Lyu, and Kamesh Munagala. Proportionally fair clustering. In *Proceedings of the 36thInternational Conference on Machine Learning (ICML)*, pages 1032–1041, 2019.

9 Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In *Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 5036–5044, 2017.

10 Vincent Conitzer, Vincent Freeman, Nisarg Shah, and Jennifer Wortman Vaughan. Group fairness for the allocation of indivisible goods. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1853–1860, 2019.

11 Joseph Cox and Michael B Partensky. Spatial localization problem and the circle of Apollonius. arXiv, 2007. `arXiv:physics/0701146`.

12 Brandon Fain, Kamesh Munagala, and Nisarg Shah. Fair allocation of indivisible public goods. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*, pages 575–592, 2018.

13 Michal Feldman and Yoav Wilf. Strategyproof facility location and the least squares objective. In *Proceedings of the fourteenth ACM conference on Electronic Commerce (EC)*, pages 873–890, 2013.

14 Dimitris Fotakis. Incremental algorithms for facility location and k-median. *Theoretical Computer Science*, 361(2-3):275–313, 2006.

15 Anupam Gupta, Guru Guruganesh, and Melanie Schmidt. Approximation algorithms for aversion k-clustering via local k-median. In *Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP)*, pages 66:1–66:13, 2016.

16 Moritz Hardt, Eric Price, and Nathan Srebro. Equality of opportunity in supervised learning. In *Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 3315–3323, 2016.

17 Tatsunori B Hashimoto, Megha Srivastava, Hongseok Namkoong, and Percy Liang. Fairness without demographics in repeated loss minimization. arXiv, 2018. `arXiv:1806.08010`.

18 Stephen T. Hedetniemi, David P. Jacobs, and K. E. Kennedy. A theorem of ore and self-stabilizing algorithms for disjoint minimal dominating sets. *Theoretical Computer Science*, 593:132–138, 2015.

19 Safwan Hossain, Andjela Mladenovic, and Nisarg Shah. Designing fairly fair classifiers via economic fairness notions. In *Proceedings of the 29th International World Wide Web Conference (WWW)*, pages 1559–1569, 2020.

20 Michael Kearns, Seth Neel, Aaron Roth, and Zhiwei Steven Wu. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In *Proceedings of the 35thInternational Conference on Machine Learning (ICML)*, pages 2569–2577, 2018.

21 Ke Liao and Diansheng Guo. A clustering-based approach to the capacitated facility location problem 1. *Transactions in GIS*, 12(3):323–339, 2008.

22 Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. arXiv, 2019. `arXiv:1908.09635`.

23 Hervé Moulin. *Fair Division and Collective Welfare*. MIT Press, 2003.

24 Arvind Narayanan. Translation tutorial: 21 fairness definitions and their politics. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT)*, 2018.

25 Cathy O'Neil. *Weapons of math destruction: How big data increases inequality and threatens democracy.* Broadway Books, 2016.

26 Oystein Ore. Theory of graphs. *American Mathematical Society Colloquium. Providence, R.I.*, 38, 1962.

27 David C. Parkes and Rakesh V. Vohra. Algorithmic and economic perspectives on fairness. arXiv, 2019. `arXiv:1909.05282`.

28   Clemens Rösner and Melanie Schmidt. Privacy preserving clustering with constraints. In *Proceedings of the 45th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 96:1—96:14, 2018.

29   Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge University Press, 2014.

30   Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, 1974.

31   Vladimir Shenmaier. The problem of a minimal ball enclosing k points. *Journal of Applied and Industrial Mathematics*, 7(3):444–448, 2013.

32   Berk Ustun, Yang Liu, and David C. Parkes. Fairness without harm: Decoupled classifiers with preference guarantees. In *Proceedings of the 36thInternational Conference on Machine Learning (ICML)*, pages 6373–6382, 2019.

33   Vladimir Vapnik and Alexey Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264—280, 1971.

34   Hal Varian. Equity, envy and efficiency. *Journal of Economic Theory*, 9:63–91, 1974.

35   Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, and Krishna P. Gummadi. Fairness constraints: A flexible approach for fair classification. *Journal of Machine Learning Research*, 20(75):1–42, 2019.

36   Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. Learning fair representations. In *Proceedings of the 30thInternational Conference on Machine Learning (ICML)*, pages 325–333, 2013.