Distribution Agreement

Zirui Deng April 6, 2021

Fair Clustering Problems

by

Zirui Deng

Michelangelo Grigni
Adviser

Department of Mathematics

Michelangelo Grigni
Adviser

Jinho Choi
Committee Member

Manuela Manetta
Committee Member

2021

Fair Clustering Problems

by

Zirui Deng

Michelangelo Grigni
Advisor

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Mathematics

2021

Abstract

Fair Clustering Problems

By Zirui Deng

Clustering is a fundamental tool in machine learning and data mining which takes many forms according to different objectives being considered. It entails partitioning points into groups (clusters) and may be used to make decisions for each point based on its group. We study various fair clustering problems under the disparate impact doctrine, where each minority class must be represented approximately equally in every cluster. We offer a survey of recent clustering algorithms that account for different notions of fairness.

Fair Clustering Problems

By

Zirui Deng

Michelangelo Grigni
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Mathematics

2021

Acknowledgements

I wholeheartedly thank Dr. Michelangelo Grigni for introducing me to the notion of fairness in machine learning and clustering in particular and offering me invaluable suggestions and feedback through this work. I would also like to thank Dr. Jinho Choi and Dr. Manuela Manetta for taking the time to serve on my thesis committee and giving me meaningful advice on my writing and presentation. Finally, I would like to thank my family for supporting my studies at Emory University and letting me choose my career path afterwards.

Table of Contents

List of Figures

1 Introduction

Machine learning algorithms and their potential bias towards underrepresented groups have become increasingly visible. While the learning algorithms themselves are not necessarily biased or unfair in nature, they may still run the unintended risk of exposing and amplifying biases that are already present in the training data available. Thus the recent work that focuses on designing fair algorithms is certainly justified. There has been a growing interest in developing research into fairness in various learning and optimization problems. The goal here is to develop reasonable criteria and algorithms to ensure that we are able to find solutions for optimization problems that are fair with respect to a certain protected feature such as gender or race.

Recent exploration of fairness in machine learning mainly involves two directions. The first is in defining what fairness means exactly. We trace the inception of individual fairness considerations to Dwork et al. [DHP+12]. In their framework, fairness is characterized by the principle that any two individuals who are similar with respect to a particular feature or task should be classified similarly. They then assume a distance metric that defines similarity between individuals. Specifically, the guiding principle of fairness is formalized as a Lipschitz condition on the classifier [DHP+12]. The Lipschitz condition requires that any two individuals x, y that are at distance $d(x, y) \in [0, 1]$ are mapped to distributions $M(x)$ and $M(y)$, respectively, with the property that the statistical distance between $M(x)$ and $M(y)$ is at most $d(x, y)$. In other words, the distributions over outcomes observed by x and y are indistinguishable up to $d(x, y)$. The optimization problem of constructing fair Lipschitz classifiers can be expressed as a linear program and solved efficiently.

The second direction, which is the focus of this work, is in designing algorithms that yield fair outcomes according to one specific notion of fairness. All of the fair algorithms we will proceed to dig into come from this line of exploration. They also share the notion of disparate impact due to Feldman et al. [FFM+15], which informally states that protected attributes such as race and gender should not explicitly affect decision making, and the decisions made should not reflect disproportionately differences for applicants in different protected classes [CKLV17].

In this work we consider the problem of clustering, one of the most common unsupervised problems. Typically with clustering problems we are given a set X of n points that are located in some metric space. We aim to find a clustering, i.e., a partition of X into different groups (or clusters) that result in an optimal value with respect to a certain objective function. In the metric k-center problem, for instance, our goal is to find a set of k vertices for which the largest distance of any point to its closest vertex in the chosen set of k vertices is minimum. In the k-median problem, our goal is to find a set of k vertices for which the sum of the distances from the points in X to their nearest centers is minimum. The vertices must be in a metric space, thus providing a complete graph that satisfies the triangle inequality.

The classical clustering problems can all be solved exactly, but in exponential time. Since we would like to avoid working with NP-hard problems, we instead turn to approximation techniques that give rise to substantially faster algorithms. There have been several proposed approximate solutions to the classical k-center and k-median problems. A greedy algorithm given by Gonzalez [Gon85] is known to be a 2-approximation for the k-center problem. Another interesting algorithm given by Hochbaum and Shmoys is also a 2-approximation [HS85]. Their method uses the technique of parameter pruning [Vaz03]. The k-median clustering problem can be approximated by a linear programming relaxation of an integer program. Assume that the points in X are indexed by $1, 2, ..., n$.

$$\min \sum_{i,j} x_{ij} d(i,j)$$

$$\sum_j y_j \leq k$$

$$\sum_j x_{ij} = 1$$

$$x_{ij} \leq y_j$$

$$x_{ij}, y_j \in \{0, 1\}$$

where $d(i,j)$ is the distance between two points i, j ($1 \leq i, j \leq n$), $y_j = 1$ iff point j is chosen as a center, and $x_{ij} = 1$ iff j is the center that serves point i. Because an integer program in general cannot be solved efficiently, we relax it into a linear program that can be solved in linear time, by replacing

$$x_{ij}, y_j \in \{0, 1\}$$

with

$$x_{ij}, y_j \in [0, 1].$$

Then we can round the (fractional) solution to this linear program to get an approximate solution to the original problem.

For convenience in the subsequent discussion, we assume that each point $x \in X$ has a color that helps identify its protected class. A clustering algorithm that does not take the notion of protected attributes into consideration is called *colorblind* [CKLV17], and it may result in unfair clusterings. Figure 1 illustrates the difference between fair and colorblind algorithms.

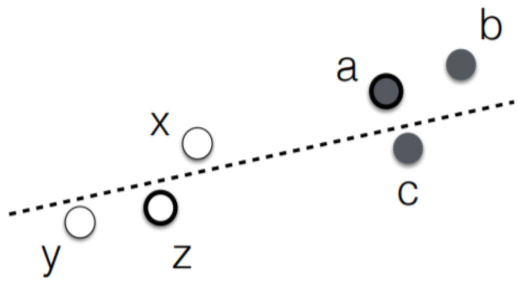

Figure 1: Illustration for the difference between fair and unfair (colorblind) algorithms.

As we know, a classic, colorblind k-center clustering algorithm would group points a, b, c into one cluster and x, y, z into another, with a and z as centers. A fair k-center clustering algorithm, on the other hand, may group a, b, x into one cluster and c, y, z into another with a and z as centers, as indicated by the dashed line, so that every protected class is represented relatively equally in each cluster. We observe that in the fair case a point is no longer necessarily assigned to its nearest cluster center; in this example, x is assigned to a even though z is closer.

One real-world problem that captures the essence of the above figure would be like this [BGK$^+$19]: Suppose we are given the task of assigning incoming students to k schools such that the maximum distance of any student to his or her nearest school is minimized. The school capacity is indicated by the number of its teachers: For each teacher, s students, for example, can be admitted. It is easy to tell that this is actually an instance of the metric k-center problem. A naïve solution may result in some schools having more boys than girls and vice versa. We would prefer a fair assignment where the classes are more balanced in terms of gender. We need to assign the children so that the gender ratio is approximately 1:1 while minimizing the maximum distance.

We will be discussing fair approximation algorithms in subsequent sections. The algorithms all run in polynomial time. In Section 2, we first examine the work by Chierichetti et al. [CKLV17] that employs the notion of *fairlets* for solving fair k-center and k-median problems. A fairlet is a small set of elements that satisfies the fairness constraint. Chierichetti et al. [CKLV17] have shown that fair k-median and k-center can be solved by first decomposing an instance into fairlets and then solving the clustering problem on the set of centers of these fairlets. Their main results are a 4-approximation for the fair k-center problem and $(t+1+\sqrt{3}+\epsilon)$-approximation for the k-median problem where t is the balance (defined in section 2) of the resulting clustering from the algorithm.

In Section 3, the idea of *restricted dominance* (RD) and *minority protection* (MP) is employed by Bera et al. [BCFN19] in their presentation of fair algorithms for clustering. Additionally, the results by Bera et al. have to do with the amount of overlap, Δ, defined to be the maximum number of protected groups any individual can be a part of: Given any ρ-approximation algorithm for a classic clustering problem, a $(\rho+2)$-approximation with $4\Delta+3$ additive error for the corresponding fair clustering problem can be achieved. In the special case of $\Delta = 1$, meaning no overlap between clusters, the additive error can be improved to 3 [BCFN19].

In Section 4, we investigate the *fair correlation clustering problem* brought forth by Ahmadian et al. [AEKM20], which takes multiple colors into account and uses information about *similarity* and *dissimilarity* relationships among a set of points. In contrast to other k-median and k-center problems, the number of clusters is not specified beforehand but rather determined based on the outcome of an optimization. Their main contribution lies in a fairlet-based reduction for the graph clustering problem of correlation clustering. They have introduced a cost function that caters to the correlation clustering fairlet decomposition problem and proven that this cost can be approximated by a median-type clustering cost function for a carefully defined metric space [AEKM20]. Their results depend on an upper bound of α on the fraction of vertices of every color of every each cluster.

In Section 5, we consider the notion of *essential fairness* introduced by Bercea et al. [BGK+19] A cluster C is said to be *essentially fair* if there exists a fractional fair cluster C' such that for each color h the number of color h points in C differ by at most 1 from the mass of color h points in C', and this difference induces a small additive fairness violation. Bercea et al. start with a solution to the unconstrained problem. and derive a fair clustering solution with the same centers. They have achieved this by a technique called *weakly supervised LP rounding* [BGK+19], in that they solve an LP for the fair clustering problem and then combine it with the integral unfair solution by careful rounding. They have obtained results in this way for a wide range of clustering problems, not limited to k-center and k-median. We also notice that the bounds used in this section are similar to the idea of *restricted dominance* (RD) and *minority protection* (MP) in Section 3.

There is some notational convention shared by the subsequent sections. For an integer k, let $[k]$ denote the set $\{1, ..., k\}$. For two points u, v in a vertex set V, (u, v) denotes the edge that connects u and v, and $d(u, v)$ denotes the distance between u and v.

2 Fair Clustering Through Fairlets

Let X be a set of points in a metric space. In this section we assume that each point in X is colored either red or blue. If R denotes the set of red points and B the set of blue points, then R and B are disjoint and $X = R \cup B$.

2.1 Definitions

We first define a notion of balance with regard to the coloring of points as follows.

Definition 1 (Equivalent to Definition 1 of [CKLV17]). *For a non-empty subset $X' \subseteq X$, the* balance *of X' is defined as:*

$$\text{balance}(X') = \min\left(\frac{|\text{R} \cap \text{X'}|}{|\text{B} \cap \text{X'}|}, \frac{|\text{B} \cap \text{X'}|}{|\text{R} \cap \text{X'}|}\right) \in [0, 1]$$

where $|\text{R} \cap \text{X'}|$ and $|\text{B} \cap \text{X'}|$ denote the number of red points (blue points, respectively) in X'. A clustering is a partition of all client points into disjoint subsets. The balance *of a clustering $C = \{C_1, C_2, ...\}$ is defined as:*

$$\text{balance}(C) = \min_{C_i \in C} \text{balance}(C_i).$$

In order to introduce the notion of fairness and fairlets employed in this section, we need to prove the following lemma.

Lemma 1 (Equivalent to Lemma 3 of [CKLV17]). *Let $b/r \leq \text{balance}(X) \leq 1$ for some integers $1 \leq b \leq r$ such that $gcd(b, r) = 1$. Then there exists a clustering $Y = \{Y_1, ..., Y_m\}$ of X such that: (i) $|Y_j| \leq b + r$ for each $Y_j \in Y$, i.e., each cluster is small, and (ii)* $\text{balance}(Y) \geq b/r$.

Proof. Without loss of generality, let $|\text{B} \cap \text{X'}| \leq |\text{R} \cap \text{X'}|$. By assumption, $\frac{|\text{B} \cap \text{X'}|}{|\text{R} \cap \text{X'}|} \geq \frac{b}{r}$.

Case 1: $\frac{|\text{B} \cap \text{X'}|}{|\text{R} \cap \text{X'}|} > \frac{b}{r}$. We construct the clustering Y iteratively as follows.

If $|\text{R} \cap \text{X'}| - |\text{B} \cap \text{X'}| \geq r - b$, then we remove r red points and b blue points from the current set to form a cluster $Y_j \in Y$, assuming that there are enough points from which to remove to form the cluster. By construction, $|Y_j| = b + r$ and $\text{balance}(Y_j) = b/r$.

Furthermore the leftover set has balance (|B ∩ X'| $-b$)/(|R ∩ X'| $-r$) $\geq b/r$ and we iterate on this leftover set.

If |R ∩ X'| − |B ∩ X'| $< r - b$, then we remove (|R ∩ X'| − |B ∩ X'| $+b$) red points and b blue points from the current set to form $Y_j \in Y$, assuming that there are enough points from which to remove to form the cluster. Note that $|Y_j| \leq b + r$ and that balance(Y_j) $= b/($|R ∩ X'| − |B ∩ X'| $+b) \geq b/r$.

Finally note that when the remaining points are such that the red and the blue points are in a one-to-one correspondence, we can pair them up into perfectly balanced clusters of size 2.

Case 2: $\frac{|\mathrm{B} \cap \mathrm{X'}|}{|\mathrm{R} \cap \mathrm{X'}|} = \frac{b}{r}$, which means |B ∩ X'| $= mb$ and |R ∩ X'| $= mr$. We can simply construct a clustering Y consisting of m clusters, each of which contains b blue points and r points. □

We call the clustering Y as described in Lemma 1 a *(b, r)-fairlet decomposition* of X and call each cluster $Y_j \in Y$ a fairlet. Now we proceed to define fair clustering problems with respect to the clustering objectives. To do so, we need to assume that the metric space containing X is equipped with a distance function $d : X^2 \to \mathbb{R}^{\geq 0}$.

Definition 2 (Equivalent to Definition 4 of [CKLV17])**.** *The problem of dividing X into a clustering C such that (i) $|C| = k$, (ii)* balance(C) $\geq t$*, and (iii)* $\phi(X, C) = \max_{C_i \in C} \min_{c \in C_i} \max_{x \in C_i} d(x, c)$ *is minimized is called the* (t, k)-fair center problem.

Similarly, for the definition of the (t, k)-fair median problem*, the goal is to divide X into C such that (i) $|C| = k$, (ii)* balance(C) $\geq t$ *and (iii)* $\psi(X, C) = \sum_{C_i \in C} \min_{c \in C_i} \sum_{x \in C_i} d(x, c)$ *is minimized.*

For the definition of the (t, k)-fair means problem*, the goal is to divide X into C*

such that (i) $|C| = k$, *(ii)* $\text{balance}(C) \geq t$ *and (iii)* $\psi(X, C) = \sum_{C_i \in C} \min_{c \in C_i} \sum_{x \in C_i} (d(x, c))^2$ *is minimized.*

Here are some extra notations we need to introduce before defining decomposition cost. Let $Y = \{Y_1, ..., Y_m\}$ be a fairlet decomposition. For each cluster Y_i, we let an arbitrary point $y_i \in Y_i$ be its center. Let $Y' = \{y_1, ..., y_m\}$ be the set of fairlet centers from Y. For any point $x \in X$, we denote $\beta : X \to [m]$ as the index of the fairlet to which it is mapped. That is, the index of the fairlet to which x belongs is $\beta(x)$. Thus the center to which x is assigned can be expressed as $y_{\beta(x)}$. We may now define the cost of a fairlet decomposition.

Definition 3 (Equivalent to Definition 5 of [CKLV17])**.** *For a fairlet decomposition, we define its* k-center cost *as* $\max_{x \in X} d(x, y_{\beta(x)})$*, its* k-median cost *as* $\sum_{x \in X} d(x, y_{\beta(x)})$*, and its* k-means cost *as* $\sum_{x \in X} (d(x, y_{\beta(x)}))^2$*. A (b, r)-fairlet decomposition is optimal if it has minimum cost among all possible (b, r)-fairlet decompositions.*

2.2 Main results

Here we present the main results of [CKLV17] that show that any fair clustering problem can be reduced to finding a fairlet decomposition through constructions of min-cost flow (MCF) instances and then clustering the resulting fairlets using classical algorithms. Given a standard graph with a source, a sink, edge costs and capacities, the focus of an MCF problem is to push a certain amount of flow from source to sink at a minimum cost while maintaining flow conservation and satisfying capacity constraints. The MCF problem can be solved in polynomial time.

Theorem 1 (Equivalent to Theorem 13 of [CKLV17])**.** *There exists a 4-approximation algorithm for the (1/t', k)-fair center problem, where* $t = 1/t'$*, for any positive integer* t'*. The algorithm first finds fairlets and then clusters them.*

Proof. Here we mainly contribute to a proof of the 2 factor omitted in [CKLV17]. The creation of a min-cost flow (MCF) instance is similar to that in [CKLV17]. Suppose $Z = \{Z_1, ..., Z_k\}$ is an optimal $(1/t', k)$-fair center clustering with cost C. That is, Z has k clusters, each with balance at least $1/t'$, and all points are within distance C of their cluster center. Our goal is to find a (1/t',k)-fair center clustering of cost at most 4C.

By applying Lemma 1 to each cluster of Z, with $b = 1$ and $r = t'$, we get a (1,t')-fairlet decomposition Y. For every point $x \in X$, we let $Z_j(x)$ denote the center of the cluster $Z_j \in Z$ containing x, and from the notations introduced earlier we have $y_{\beta(x)}$ as the center of the cluster $Y_{\beta(x)} \in Y$ containing x. The maximum distance between any two points in one fairlet is at most 2C (In particular, all red-blue distances are at most 2C): Since Z is optimal, the distance between x and $Z_j(x)$, $d(x, Z_j(x))$, and the distance between $y_{\beta(x)}$ and $Z_j(x)$, $d(y_{\beta(x)}, Z_j(x))$, are both at most C. Thus by triangular inequality, $d(x, y_{\beta(x)}) \leq d(x, Z_j(x)) + d(y_{\beta(x)}, Z_j(x)) = 2\text{C}$.

Now we need to set up an appropriate MCF instance as follows. Note that for k-center, the objective is usually encoded in the constraints, and the linear program itself has no objective function. We are simply aiming for an optimal value, which we will call τ. Given the set of blue points B, the set of red points R, any integer t', and τ, which serves as a positive parameter of the algorithm, we construct a directed graph $H = (V, E)$, whose node set V consists of source node β, sink node ρ, all nodes in $B \cup R$, and t' additional copies of each node $v \in B \cup R$, i.e.

$$V = \{\beta, \rho\} \cup B \cup R \cup \{b_i^j : b_i \in B, j \in [t']\} \cup \{r_i^j : r_i \in R, j \in [t']\}$$

where b_i^j's and r_i^j's are the additional copies. The directed edges of H are made up of the following:

(i) A (β, ρ) edge with cost 0 and capacity $\min(|B|, |R|)$,

(ii) a (β, b_i) edge for each $b_i \in B$ and an (r_i, ρ) edge for each $r_i \in R$, all of whom have cost 0 and capacity $t' - 1$,

(iii) a (b_i, b_i^j) edge for each $b_i \in B$ and for each $j \in [t']$, and a (r_i, r_i^j) edge for each $r_i \in R$ and for each $j \in [t']$, all of whom have cost 0 and capacity 1,

(iv) a (b_i^m, r_j^n) edge with capacity 1 for each $b_i \in B, r_j \in R$ and for each $m, n \in [t']$, whose cost is 1 if $d(b_i, r_j) \leq \tau$ and ∞ otherwise.

Furthermore, we specify that each node in B has a supply of 1, each node in R has a demand of 1, β has a supply of $|R|$, and ρ has a demand of $|B|$. The other nodes each have zero supply and zero demand. Below we show an example of such construction for $t' = 2$. Here the only nodes with positive demands or supplies are $\beta, \rho, b_1, b_2, b_3, r_1$ and r_2, and all dotted edges have cost 0. For an example of this construction, see Figure 2.

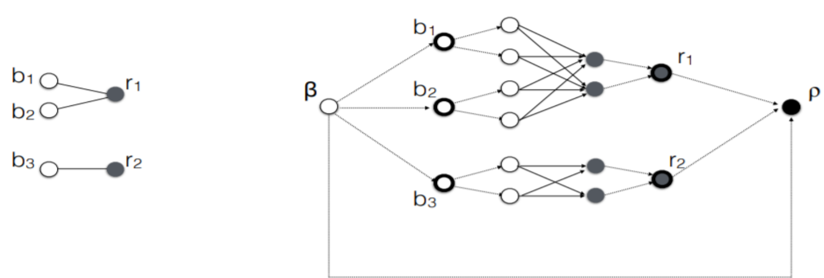

Figure 2: Example construction of an MCF instance for the directed graph where $t' = 2$.

Since all demands and capacities are integers, there exists an integral solution that pushes integral flow across each edge. If we set up the MCF instance in the manner shown above, allowing only red-blue edges of length at most 2C, we will get a (1,t')-fairlet decomposition Y, again with all red-blue distances at most 2C. (Each fairlet has either a single red point or a single blue point, which we can treat as the center of that fairlet.)

Since Y' as the set of fairlets centers from Y is a subset of X, we know Y' has a k-center decomposition with cost at most C. So the Gonzalez algorithm [Gon85] will find a k-center decomposition Z' of Y' with cost at most 2C. Now we combine Y and Z'. That is, starting with clustering Z': for each point x, add it to the same cluster as its fairlet center $y_{\beta(x)}$. Then x is at distance at most 2C from $y_{\beta(x)}$, and $y_{\beta(x)}$ is at distance at most 2C from its cluster center in Z'. Thus we have a (1/t',k)-fair center clustering of cost at most 2C + 2C = 4C. □

Theorem 2 (Equivalent to Theorem 15 of [CKLV17])**.** *There exists a* $(t'+1+\sqrt{3}+\epsilon)$*-approximation algorithm for the (1/t', k)-fair median problem for any positive integer* t'*. The algorithm first finds fairlets and then clusters them.*

Proof. Here we mainly contribute to a proof of the t' factor omitted in [CKLV17]. The creation of an MCF instance is similar to that in [CKLV17]. Suppose $Z = \{Z_1, ..., Z_k\}$ is an optimal $(1/t', k)$-fair median clustering with cost C. That is, Z has k clusters, each with balance at least $1/t'$, and the sum of the distances of every point to its cluster center is at most C. Our goal is to find a (1/t',k)-fair median clustering of cost at most $(t' + 1 + \sqrt{3} + \epsilon)$C.

By applying Lemma 1 to each cluster of Z, with $b = 1$ and $r = t'$, we get a (1,t')-fairlet decomposition Y. For every point $x \in X$, we let $Z_j(x)$ denote the center of the cluster $Z_j \in Z$ containing x. By triangular inequality, $d(x, y_{\beta(x)}) \leq d(x, Z_j(x)) + d(y_{\beta(x)}, Z_j(x))$. Then since Z is optimal, we take the sum over all points in X and get:

$$\begin{aligned}\sum_{x\in X} d(x, y_{\beta(x)}) &\leq \sum_{x\in X} d(x, Z_j(x)) + (t'-1)\sum_{x\in Y'} d(x, Z_j(x)) \\ &\leq \sum_{x\in X} d(x, Z_j(x)) + (t'-1)\sum_{x\in X} d(x, Z_j(x)) \\ &= t'\sum_{x\in X} d(x, Z_j(x)) \\ &= t'C.\end{aligned}$$

We create an MCF instance like we do with the k-center problem with the exception that we set the cost of every red-blue edge to the distance between the corresponding red node and blue node. That is, for each $b_i \in B, r_j \in R$ and for each $m, n \in [t']$, we set the cost of the edge (b_i^m, r_j^n) to $d(b_i, r_j)$. In this way we will get a (1,t')-fairlet decomposition Y with the sum of all red-blue distances at most t'C. (Each fairlet has either a single red point or a single blue point, which we can treat as the center of that fairlet.)

Since Y' as the set of fairlets centers from Y is a subset of X, we know Y' has a k-median decomposition with cost at most C. So the Li & Svensson algorithm [LS16] will find a k-median decomposition Z' of Y' with cost at most $(1+\sqrt{3}+\epsilon)$C. Now we combine Y and Z' in a fashion similar to that in Theorem 1 and get a (1/t',k)-fair median clustering of cost at most t'C + $(1+\sqrt{3}+\epsilon)$C = $(t'+1+\sqrt{3}+\epsilon)$C. □

Theorem 3. *There exists a* $(2t' + 6.357)$*-approximation algorithm for the (1/t', k)-fair means problem for any positive integer* t'*. The algorithm first finds fairlets and then clusters them.*

Proof. Here we mainly contribute to a proof of the $2t'$ factor. The creation of an MCF instance is similar to that in [CKLV17]. Suppose $Z = \{Z_1, ..., Z_k\}$ is an optimal $(1/t'$,

k)-fair means clustering with cost C. That is, Z has k clusters, each with balance at least $1/t'$, and the sum of the distances squared of every point to its cluster center is at most C. Our goal is to find a (1/t',k)-fair means clustering of cost at most $(2t' + 6.357)$C.

By applying Lemma 1 to each cluster of Z, with $b = 1$ and $r = t'$, we get a (1,t')-fairlet decomposition Y. For every point $x \in X$, we let $Z_j(x)$ denote the center of the cluster $Z_j \in Z$ containing x. By triangular inequality, $d(x, y_{\beta(x)}) \leq d(x, Z_j(x)) + d(y_{\beta(x)}, Z_j(x))$. Squaring both sides of this we get another inequality:

$$\begin{aligned}(d(x, y_{\beta(x)}))^2 &\leq (d(x, Z_j(x)) + d(y_{\beta(x)}, Z_j(x)))^2 \\ &= (d(x, Z_j(x)))^2 + (d(y_{\beta(x)}, Z_j(x)))^2 + 2d(x, Z_j(x))d(y_{\beta(x)}, Z_j(x)) \\ &\leq (d(x, Z_j(x)))^2 + (d(y_{\beta(x)}, Z_j(x)))^2 + (d(x, Z_j(x)))^2 + (d(y_{\beta(x)}, Z_j(x)))^2 \\ &= 2(d(x, Z_j(x)))^2 + 2(d(y_{\beta(x)}, Z_j(x)))^2\end{aligned}$$

Then since Z is optimal, we take the sum over all points in X and get:

$$\begin{aligned}\sum_{x \in X}(d(x, y_{\beta(x)}))^2 &\leq 2\sum_{x \in X}(d(x, Z_j(x)))^2 + 2(t'-1)\sum_{x \in Y'}(d(x, Z_j(x)))^2 \\ &\leq 2\sum_{x \in X}(d(x, Z_j(x)))^2 + 2(t'-1)\sum_{x \in X}(d(x, Z_j(x)))^2 \\ &= 2t'\sum_{x \in X}(d(x, Z_j(x)))^2 \\ &= 2t'C.\end{aligned}$$

We create an MCF instance like we do with the k-center problem with the exception that we set the cost of every red-blue edge to the distance squared between the corresponding red node and blue node. That is, for each $b_i \in B, r_j \in R$ and for each $m, n \in [t']$, we set the cost of the edge (b_i^m, r_j^n) to $(d(b_i, r_j))^2$. In this way we will get a

(1,t')-fairlet decomposition Y with the sum of all red-blue distances squared at most $2t'$C. (Each fairlet has either a single red point or a single blue point, which we can treat as the center of that fairlet.)

Since Y' as the set of fairlets centers from Y is a subset of X, we know Y' has a k-means decomposition with cost at most C. The best known approximation algorithm for classical k-means will find a k-means decomposition Z' of Y' with cost at most 6.357C [ANFSW17]. Now we combine Y and Z' in a fashion similar to that in Theorem 1 and get a (1/t',k)-fair means clustering of cost at most $2t'\text{C} + 6.357\text{C} = (2t' + 6.357)\text{C}$. □

Remark: The results of **Theorem 1** and **Theorem 2** have been claimed in [CKLV17]. **Theorem 3** is a completely new result that has not been present in existing research literature in the fair clustering area, until now.

3 Algorithms for "Relaxed" Fair Clustering

Building on the discussion in the previous section, we move on to consider the more general situation where individuals are allowed to belong to more than one protected group. We denote Δ to be the amount of overlap, i.e. the maximum number of groups of which any individual can be a member. For example, if X is the set of clients and $X_1, X_2, ..., X_l$ are l groups of X with $X = \bigcup_{i\in[l]} X_i$, then the X_i's are disjoint if and only if $\Delta = 1$. The quality of solutions to the fair clustering problems depends on Δ. To obtain more fair algorithms for a wider variety of problems, we consider a somewhat "relaxed" version of fairness where we allow very small additive violations to the fairness constraint. We are able to get fair algorithms for any l_p-norm objective if small additive violations are allowed, and these violations are shown to be negligible through empirical studies on large data sets [BCFN19].

For each group i, we define two parameters $\beta_i, \alpha_i \in [0, 1]$. We say that a clustering solution is fair if each cluster satisfies restricted dominance (RD), which states that the fraction of individuals from group i in any cluster is at most α_i, and minority protection (MP), which asserts that the fraction of individuals from group i in any cluster is at least β_i [BCFN19]. β_i, α_i's can be arbitrary parameters. With this notion, we are able to define fair clustering problems in this section using (RD) and (MP) fairness constraints.

We present a two-step procedure for solving the fair clustering problem. First, we solve the classical clustering problem and fix the centers S. Then, we solve a fair assignment problem, which we proceed to define below, on the same set of facilities to get assignment ϕ. We return (S, ϕ) as our fair solution.

Definition 4 (Equivalent to Definition 2 of [BCFN19])**.** *Given the original set of clients X and l groups of X, $X_1, X_2, ..., X_l$, with $X = \bigcup_{i\in[l]} X_i$, and a set $S \subseteq F$ with $|S| = k$, where F is the set of all possible cluster center locations, the objective of the fair assignment problem is to find the* assignment $\phi : X \to S$ *such that (a) the following* (RD) *and* (MP) *constraints are satisfied:*

$$|\{v \in X_i : \phi(v) = f\}| \leq \alpha_i \cdot |\{v \in X : \phi(v) = f\}|, \forall f \in S, \forall i \in [l] \tag{RD}$$

$$|\{v \in X_i : \phi(v) = f\}| \geq \beta_i \cdot |\{v \in X : \phi(v) = f\}|, \forall f \in S, \forall i \in [l], \tag{MP}$$

and (b) the objective function $L(S; \phi) = (\Sigma_{v\in X} d(v, \phi(v))^p)^{1/p}$ is minimized among all such assignments. Here, $p = 1, 2, \infty$ correspond to the fair k-median, k-means, and k-center problems respectively.

We denote as $OPT_v(I)$ the optimal value of the classical clustering problem (like those defined in the introduction section) given an instance I of the problem. Similarly, we denote as $OPT_f(I)$ the optimal value of any instance I of the fair clustering problem and $OPT_a(I)$ the optimal value of any instance I of the assignment problem defined above. Before we show a reduction from the fair clustering problem to the fair assignment problem, we need to introduce the notion of λ-additive violation:

A fair solution (S, ϕ) has λ-additive violation if and only if the (RD) and (MP) constraints are satisfied with a margin of error of λ. More formally, for any $f \in S$ and for any $i \in [l]$, we have

$$\beta_i \cdot |\{v \in X : \phi(v) = f\}| - \lambda \leq |\{v \in X_i : \phi(v) = f\}| \leq \alpha_i \cdot |\{v \in X : \phi(v) = f\}| + \lambda.$$

Theorem 4 (Equivalent to Theorem 3 of [BCFN19]). *Given a ρ-approximation algorithm A for the classical clustering problem and an algorithm B with λ-additive violation for the fair assignment problem, there is a $(\rho + 2)$-approximation algorithm for the fair clustering problem with λ-additive violation.*

Proof. Given an instance I of the fair clustering problem, we run algorithm A on I to get a solution (S, ϕ) (not necessarily fair). We have $L(S; \phi) \leq \rho \cdot OPT_v(I) \leq \rho \cdot OPT_f(I)$. Let J be the instance of the fair assignment problem obtained by taking S as the set of facilities. We run B on J to get an assignment $\hat{\phi}$ with λ-additive violation. $(S, \hat{\phi})$ is the desired solution. The rest of the proof follows from the lemma below. $\square$

Lemma 2 (Equivalent to Lemma 4 of [BCFN19]). *$OPT_a(J) \leq (\rho + 2) \cdot OPT_f(I)$.*

Proof. Suppose the optimal solution of I is (S^*, ϕ^*) with $L(S^*; \phi^*) = OPT_f(I)$. Recalling that (S, ϕ) is the solution returned by algorithm A, we describe the existence of an assignment $\phi' : X \to S$ such that ϕ' satisfies (RD) and (MP) constraints and

$L(S;\phi') \le (\rho+2)\cdot OPT_f(I)$: For every $f^* \in S^*$, define $\mathrm{nrst}(f^*) := argmin_{f\in S} d(f, f^*)$ as the closest facility in S to f^*. For every client $c \in X$, define $\phi'(c) := \mathrm{nrst}(\phi^*(c))$. It is rather straightforward to prove that (RD) and (MP) constraints are satisfied. Please refer to [BCFN19] for details.

Now we proceed to show $L(S;\phi') \le (\rho+2)\cdot OPT_f(I)$: Fix a client $c \in X$. Let $f = \phi(c), f' = \phi'(c), f^* = \phi^*(c)$. With the definition of nrst and triangular inequality, we have

$$d(c, f') = d(c, \mathrm{nrst}(f^*)) \le d(v, f^*) + d(f^*, \mathrm{nrst}(f^*)) \le d(v, f^*) + d(f^*, f) \le 2d(v, f^*) + d(v, f).$$

Then it is easy to verify that

$$L(S;\phi') \le 2L(S^*;\phi^*) + L(S;\phi) \le 2OPT_f(I) + \rho \cdot OPT_f(I) = (\rho+2)\cdot OPT_f(I).$$

$\square$

We now present a theorem that helps us ultimately achieve our main result. The proof of this theorem depends on an algorithm for the minimum degree-bounded matroid basis problem [KLS12]. For this problem, we are typically given a matroid (V, I), a cost function $c : V \to R$, and a hypergraph (V, E). For each hyperedge $e \in E(H)$, we are also given lower and upper bounds, $l(e)$ and $u(e)$, respectively. The goal is to find a basis B such that $l(e) \le |B \cap e| \le u(e)$ for every e, with minimum cost.

Theorem 5 (Equivalent to Theorem 7 of [BCFN19])**.** *There exists an algorithm with* $(4\Delta+3)$-*additive violation for the fair assignment problem.*

Please refer to [BCFN19] for details of the proof for this result. Combining the two above theorems in this section, we arrive at the main result:

Theorem 6 (Equivalent to Theorem 1 of [BCFN19]). *Given a ρ-approximation algorithm A for the classical clustering problem, we can derive a $(\rho+2)$-approximation solution (S,ϕ) with $(4\Delta+3)$-additive violation for the fair clustering problem.*

4 Fair Correlation Clustering

The problem of correlation clustering makes use of both similarity and dissimilarity relationships within a set of objects to cluster them [BBC04]. In this problem, the number of clusters is not specified beforehand but determined by the outcome of an optimization, unlike other clustering problems like k-median and k-center.

The technique developed in Section 1 has been used only for metric space clustering problems such as k-center and k-median [AEKM20]. Here in this section we aim at developing a reduction from the graph clustering problem of correlation clustering based on the notion of fairlets. To be more specific, we derive a cost function for dealing with the correlation clustering fairlet decomposition problem and then prove that this cost can be approximated by a median-type cost function for a carefully defined metric space [AEKM20]. Then, with a solution to this problem, we show that the fair correlation clustering instance can be reduced to a regular correlation clustering instance through a graph transformation.

4.1 Problem statement

Let $G=(V,E)$ be a complete undirected graph on $|V|=n$ vertices and $\sigma : E \to R$ be a function that assigns a label to each edge. The label $\sigma(e)$ for each e is either $+1$ (indicating that the two endpoints of e are similar) or -1 (indicating that they are dissimilar). This is an instance of an unweighted version of the problem. Let

$E^+ = \{e \in E : \sigma(e) > 0\}$ be the set of positive edges and $E^- = E \setminus E^+$ be the set of non-positive edges. For subsets $S, T \subseteq V$, let $E(S) = E \cap S^2$ denote the edges inside S and $E(S,T) = E \cap (S \times T)$ denote the edges between S and T. Let $E^+(S,T) = E^+ \cap E(S,T)$ and $E^-(S,T) = E^- \cap E(S,T)$.

As established before, a clustering is a partitioning $\mathbf{C} = \{C_1, C_2, ...\}$ of V into disjoint subsets. We define *intra-cluster* and *inter-cluster* edges in a clustering $\mathbf{C}$ as $intra(\mathbf{C}) = \bigcup_{C_i \in \mathbf{C}} E(C_i)$ and $inter(\mathbf{C}) = E \setminus intra(\mathbf{C})$. The correlation clustering cost of $\mathbf{C}$ is defined as:

$$COST(G, \mathbf{C}) = \sum_{e \in intra(\mathbf{C}) \cap E^-} |\sigma(e)| + \sum_{e \in inter(\mathbf{C}) \cap E^+} |\sigma(e)| = |intra(\mathbf{C}) \cap E^-| + |inter(\mathbf{C}) \cap E^+|.$$

Our goal is to find a clustering $\mathbf{C}$ that minimizes $COST(G, \mathbf{C})$ and also satisfies fairness constraints. We give a general reduction from fair correlation clustering to a median fairlet decomposition that works for any notion of fairness that has been considered.

4.2 Overview of results

Consider a fairlet decomposition $P = \{P_1, P_2, ...\}$. For each fairlet P_i, we let $FCOST^{in}(P_i) = |E^- \cap intra(P_i)|$ be the number of negative edges inside P_i. For fairlets P_i and P_j $(i \neq j)$, we let $FCOST^{out}(P_i, P_j)$ be the number of edges between them with the minority sign, i.e.,

$$FCOST^{out}(P_i, P_j) = min(|E^-(P_i, P_j)|, |E^+(P_i, P_j)|).$$

Then we let $FCOST^{in}(P) = \sum_i FCOST^{in}(P_i)$, $FCOST^{out}(P) = \sum_{i<j} FCOST^{out}(P_i, P_j)$,

and finally, $FCOST(P) = FCOST^{in}(P) + FCOST^{out}(P)$.

Given a constrained correlation clustering instance G and a fairlet decomposition P for G, we define a reduced correlation clustering instance as follows. Let G^P be a complete graph on $\{p_1, ..., p_n\}$, with $n = |P|$, where each vertex p_i corresponds to a fairlet $P_i \in P$. The label $\sigma(p_i, p_j)$ of the edge between p_i and p_j is the majority sign of the edges in $E(P_i, P_j)$ (with ties broken arbitrarily) multiplied by a weight that is equal to the number of edges in $E(P_i, P_j)$ with the majority sign. This instance G^P is a weighted instance, but as we will see, the weights of the edges are within a constant factor of each other, so we can still use unweighted correlation clustering methods. Given a solution to this unconstrained problem, we can then expand it into a solution **C'** to the original constrained problem. See **Algorithm 1**.

Algorithm 1 Constrained Correlation Clustering

1. $P \leftarrow$ approximate fairlet decomposition.
2. G^P : (p_i, p_j) gets majority sign in $E(P_i, P_j)$ and weight $\max(|E^+(P_i, P_j), |E^-(P_i, P_j)|)$.
3. Let **C** be an approximate (non-constrained) correlation clustering solution of G^P.
4. Output clustering $\mathbf{C'} = \{\cup_{p_j \in C_i} P_j : C_i \in \mathbf{C}\}$.

The first two lemmas below prove a transformation between a solution of G and a solution of G^P. The third lemma bounds the cost of a fairlet decomposition in terms of the cost of the optimal solution to the constrained correlation clustering problem. Please refer to the appendix of [BCFN19] for proof of these lemmas.

Lemma 3 (Equivalent to Lemma 3.1 of [AEKM20])**.** *Given a correlation clustering instance G, a fairlet decomposition P for G, and a clustering **C** of G, there exists a clustering **C'** of G^P such that*

$$COST(G^P, \boldsymbol{C'}) \leq COST(G, \boldsymbol{C}) + FCOST^{out}(P).$$

Lemma 4 (Equivalent to Lemma 3.2 of [AEKM20])**.** *Let **C** be a clustering of G^P*

*and **C'** be the clustering computed in **Algorithm 1**. Then,*

$$COST(G, \boldsymbol{C'}) \leq COST(G^P, \boldsymbol{C}) + FCOST(P).$$

Lemma 5 (Equivalent to Lemma 3.3 of [AEKM20])**.** *For any constrained correlation clustering instance G and any constrained clustering **C** of G, there is a fairlet decomposition P of G satisfying $FCOST(P) \leq COST(G, C)$.*

With these three lemmas we have the following:

Theorem 7 (Equivalent to Theorem 3.4 of [AEKM20])**.** *Assume there is an η-approximation algorithm A1 for finding the minimum cost fairlet decomposition P and a β-approximation algorithm A2 for solving the unconstrained correlation clustering instance G^P. Then **Algorithm 1** produces a $(\beta(1+\eta)+\eta)$-approximation for the constrained correlation clustering instance G.*

Proof. Let OPT be an optimal solution to the constrained correlation clustering instance G. By Lemma 5, the fairlet decomposition problem has a solution of cost at most COST(G, OPT). Therefore, algorithm A1 must find a decomposition P with $FCOST(P) \leq \eta \cdot COST(G, OPT)$. Also, by Lemma 3, the instance G^P has a solution of cost at most $(1 + \eta) \cdot COST(G, OPT)$. Thus algorithm A2 can find a clustering **C** in G^P of cost at most $\beta(1 + \eta) \cdot COST(G, OPT)$. Thus, by Lemma 4, the cost of the final clustering produced in G by **Algorithm 1** is at most $(\beta(1 + \eta) + \eta) \cdot COST(G, OPT)$. □

For the next step, we explain the approximation factor β we can get for solving unconstrained correlation clustering instance G^P and in the next subsection we will address approximation ratios η that we can get for minimum cost fairlet decomposition problems depending on the fairness parameter α and the number of colors in a given fair correlation clustering instance.

Lemma 6 (Equivalent to Lemma 3.5 of [AEKM20])**.** *There exists an approximation algorithm for unconstrained correlation clustering of G^P with approximation ratio of $\beta = min(log(n), 2\rho r^2)$ where $r = \frac{\max_{P_i \in P} |P_i|}{\min_{P_i \in P} |P_i|}$ and ρ is the approximation factor of unweighted correlation clustering. (Currently best known approximation factor is* $\rho = 2.06$ *[CMSY15].)*

Proof. Since the reduced correlation clustering instance is a weighted correlation clustering instance, there exists an $O(log(n))$-approximation [DEFI06]. Now since the weight of the edge between p_i and p_j in G^P is at least$|P_i| \cdot |P_j|/2$ and at most$|P_i| \cdot |P_j|$, any two edges weights are within $2r^2$ of each other. Hence if we solve the resulting unweighted instance, we will get a $(2\rho r^2)$-approximation. □

4.3 Fair decomposition

Consider a correlation clustering instance G and let d be a distance function defined on a metric space M containing the set of vertices V. For a fairlet decomposition $P = \{P_1, P_2, ...\}$, we define the following median cost:

$$MCOST(P_i) = \min_{u \in M} \sum_{v \in P_i} d(u, v)$$

and

$$MCOST(P) = \sum_{P_i \in P} MCOST(P_i).$$

Notice that the problem of finding the fairlet decomposition with minimum MCOST(P) is precisely the fairlet decomposition problem for fair k-median.

We define an embedding $\phi : V \to [0, 1]^n$ as follows. For vertices $u, v \in V$, let $\phi(u)_v = 1$

if $u = v$ or $(u, v) \in E^+$, and $\phi(u)_v = 0$ if $(u, v) \in E^-$. We then use Hamming distance as our metric. That is, for $u, v \in V$, we have $d(u, v) = |\phi(u) - \phi(v)|$.

The following two lemmas, whose proofs can be seen in the appendix of [BCFN19], show that the FCOST of a fairlet decomposition is close to its MCOST with respect to d.

Lemma 7 (Equivalent to Lemma 4.1 of [AEKM20])**.** *For any fairlet decomposition P, we have*

$$MCOST(P) \leq 2 \cdot FCOST(P).$$

Lemma 8 (Equivalent to Lemma 4.2 of [AEKM20])**.** *For any fairlet decomposition P, let $f = max_{P_i \in P}|P_i|$. Then we have*

$$FCOST(P) \leq 2f \cdot MCOST(P).$$

Hence we get the following theorem:

Theorem 8 (Equivalent to Theorem 4.3 of [AEKM20])**.** *If there is a γ-approximation algorithm for fairlet decomposition with median costs, and this algorithm always produces fairlets of size at most f, then the solution produced by this algorithm is a $(4f\gamma)$-approximation to the problem of finding a fairlet decomposition with minimum FCOST.*

Next, we focus on three fairness constraints: an upper bound of $\alpha = 1/2$ on the fraction of vertices of each color in each cluster; an upper bound of $\alpha = 1/col$ where col is the number of distinct colors, which marks an improvement over Section 2; and an upper bound of $\alpha = 1/t$ for an integer t on the fraction of vertices of each color in each cluster. (Here when we speak of the cost of a fairlet decomposition, we mean its median cost.)

4.3.1 $\alpha = 1/2$:

In this case, we can show that fairlets have size at most 3 and find these fairlets by solving a minimum weight 2-factor problem in a graph. A 2-factor is a subgraph where each vertex has degree 2 and edges may be used multiple times. Define a graph H on points in V as follows: two vertices u, v are connected by an edge if they have distinct colors, and the weight of the edge is $d(u, v)$.

Lemma 9 (Equivalent to Lemma 4.4 of [AEKM20]). *The cost of an optimal 2-factor in H can be bounded by $2 \cdot MCOST(P^*)$, where P^* is the optimal fairlet decomposition.*

Proof. We construct a feasible 2-factor by constructing a 2-factor for each fairlet $P_i \in P^*$ with center μ_i. There are at most $|P_i|/2$ vertices of any color, depending on the parity of P_i, vertices of P_i can be covered by matching and a possible multi-color triangle. Doubling the matching edges, we can get a 2-factor for covering P^*.

Then we bound the cost of this 2-factor. For a matching edge $(u, v) : u, v \in P_i$, by triangle inequality, $d(u, v) \leq d(u, \mu_i) + d(v, \mu_i)$, and for a triangle $(u, v, w) : u, v, w \in P_i$, the sum of pairwise distances can be bounded by $2(d(u, \mu_i) + d(v, \mu_i) + d(w, \mu_i))$. Therefore the cost of the proposed 2-factor for covering P^* is at most $2 \cdot MCOST(P^*)$. □

Lemma 10 (Equivalent to Lemma 4.5 of [AEKM20]). *For $\alpha = 1/2$, there is an approximation algorithm for fairlet decomposition that returns a solution with median cost at most $2 \cdot MCOST(P^*)$, the size of largest fairlet at most 3 and the size of smallest fairlet at least 2.*

Proof. Consider an optimal 2-factor in H. Define a fairlet decomposition as follows. For each cycle of even length, consider a set of alternating edges and let each alternating edge be a fairlet with one of the endpoints chosen as center. For a cycle of odd length, there must exist three consecutive vertices that have pairwise distinct colors. In this case, let one fairlet be these three vertices with the middle vertex as center and for the (unique) alternating edges covering the remaining vertices, let each edge be a fairlet with one of the endpoints as center. The median cost of these fairlets is at most the weight of the original 2-factor, which is at most $2 \cdot MCOST(P^*)$ by Lemma 9. The lemma follows. □

Lemma 10 and Theorem 5 yield the following.

Theorem 9 (Equivalent to Theorem 4.6 of [AEKM20])**.** *For $\alpha = 1/2$, there is a 256-approximation algorithm for fair correlation clustering.*

Proof. Going back to **Lemma 6**, here we take ρ to be the currently best known approximation factor 2.06. From **Lemma 10**, $r = 3/2$. Then we go back to **Theorem 8** with $f = 3$ and $\gamma = 2$. Finally we refer to **Theorem 7** with $\beta = 2\rho r^2 = 9.27$ and $\eta = 4f\gamma = 24$ to approximately reach the 256 factor. □

Remark: For the 2-color special case, vertices of P can be covered by matching alone. Then from the proof of **Lemma 8** we know that there is an approximation algorithm for fairlet decomposition that returns a solution with median cost at most $2 \cdot MCOST(P^*)$ (since edges are used twice with each matching), and the size of the fairlets are 2. Thus in this special case we have $r = 1$, $f = 2$ and $\gamma = 2$. $\beta = 2\rho r^2 = 4.12$ and $\eta = 4f\gamma = 16$. Hence there is a $(\beta(1 + \eta) + \eta) = 86$-approximation algorithm for fair correlation clustering with 2 colors.

The next case concerns multiple colors. Let *col* denote the number of distinct colors.

4.3.2 $\alpha = 1/col$:

Lemma 11 (Equivalent to Lemma 4.7 of [AEKM20]). *For $\alpha = 1/col$, there is an approximation algorithm for fairlet decomposition that returns a solution with median cost at most $col \cdot MCOST(P^*)$ and the size of each fairlet is col.*

Proof. Consider an arbitrary ordering of the colors and solve a min-cost matching problem between points of color c and $c+1$ in the graph H. The union of these matchings yields a partition of V into paths of length col. Each such path is a fairlet. Let P denote this fairlet decomposition.

Now we want to bound the cost of P. Let M be an arbitrary matching between vertices of color c and $c+1$ such that point u is matched to a point v only if u and v belong to the same partition of P^*. Since each cluster in P^* has an equal number of vertices of each color and there is an edge between any two vertices of different colors in H, a matching M exists. Since $d(u, v)$ can be bounded by $d(u, \mu) + d(v, \mu)$ where μ is the center of the partition containing u and v, the cost of M can be bounded by the median cost of serving clients of colors c and $c+1$. Since each color is matched twice, the total cost of each path corresponding to a partition is at most $2 \cdot MCOST(P^*)$. For each path we pick the middle vertex as center and the cost of assigning vertices of the path to the center is at most $col/2$, as each edge is charged at most $col/2$ times. Hence $MCOST(P) \leq col \cdot MCOST(P^*)$. This means that in this case, we get a 2-approximation with fairlets of size at most col. $\square$

Lemma 11 and Lemma 5 yield the following:

Theorem 10 (Equivalent to Theorem 4.8 of [AEKM20]). *For $\alpha = 1/col$, there is a $(20.48col^2)$-approximation algorithm for fair correlation clustering.*

Proof. The proof is similar to that of **Theorem 9**. Here we take $\rho = 2.06$ and $r = 1$. Then we go back to **Theorem 8** with $f = col$ and $\gamma = col$. Finally we refer to **Theorem 7** with $\beta = 2\rho r^2 = 4.12$ and $\eta = 4f\gamma = 4col^2$ to reach the $20.48col^2$ factor. □

Remark: The original paper claims a better bound of $16.48col^2$ [AEKM20], but we have not been able to verify it. We tried to contact the authors for an explanation, but with no response.

4.3.3 $\alpha = 1/t$:

Here t is any given integer. We argue that we can utilize any approximation algorithm for fairlet decomposition as a black-box to build an algorithm for fair correlation clustering. While we allow the black-box to produce fairlets of arbitrary size, the following lemma ensures that we are able to bound the size of the fairlets.

Lemma 12 (Equivalent to Lemma 4.9 of [AEKM20])**.** *For any set P that satisfies fairness constraint with $\alpha = 1/t$, there exists a partition of P into sets $(P_1, P_2, ...)$ where each P_i satisfies the fairness constraint and $t \leq |P_i| < 2t$.*

Proof. Let $p = mt + r$ with $0 \leq r < t$, where $p = |P|$. The fairness constraint ensures that there are at most m elements of each color. Consider the partitioning obtained as follows: Given an ordering of elements where all points of the same color are in consecutive places, assign points to sets $P_1, ..., P_m$ in a round-robin fashion. Then each set P_i gets assigned at least t elements and at most $t + r < 2t$ elements. Because there are at most m elements of each color, each set gets at most one point of any color. Hence all sets satisfy the fairness constraint as $1/|P_i| \leq 1/t$. □

Theorem 11 (Equivalent to Theorem 4.10 of [AEKM20])**.** *For $\alpha = 1/t$, given an γ-approximation algorithm for fairlet decomposition with median cost, there is an $O(t\gamma)$-approximation algorithm for fair correlation clustering.*

5 Essentially Fair Clustering

Given a set of points X, in this section we also consider the case of multiple colors. Specifically, we are given a set of colors $Col = \{col_1, ..., col_g\}$ and a coloring function $col : X \to Col$ that assigns a color to each point $x \in X$. For any subset of points $X' \subseteq X$ and any color $col_h \in Col$, we denote $col_h(X') = \{h \in X' : col(h) = col_h\}$ as the set of points in X' that have color col_h and $r_h(X') = \frac{|col_j(X')|}{|X'|}$ as the ratio of col_h in X'.

In our clusters, we want to preserve the ratios of colors in X. Here we distinguish two cases: *exact* preservation of ratios and *relaxed* preservation of ratios [BGK$^+$19]. For the *exact* preservation of ratios, we ask that all clusters be *exactly fair* [BGK$^+$19]: A set of points $X' \subseteq X$ is *exactly fair* if for each $col_h \in Col$ we have $r_h(X') = r_h(X)$. For the *relaxed* case, we are given upper and lower bounds $l = (l_1 = p_1^1/q_1^1, ..., l_g = p_1^g/q_1^g)$ and $u = (u_1 = p_2^1/q_2^1, ..., u_g = p_2^g/q_2^g)$ on the ratio of colors in each cluster and require that all clusters satisfy $r_h(X') \in [l_h, u_h]$ for each color $col_h \in Col$.

5.1 LP formulations for fair clustering problems

Let $I = (X, L, col, d, k, l, u)$ be a problem instance for a fair clustering problem, where L is the set of potential locations. Let $S \subseteq L$ denote the set of locations that are opened. We introduce binary variables $y_i \in \{0, 1\}$ for all $i \in L$ such that $y_i = 1$ if and only if $i \in S$. Similarly, we introduce binary variables $x_{ij} \in \{0, 1\}$ for all $i \in L$

and $j \in P$ with $x_{ij} = 1$ if and only if j is assigned to i. All integer LP formulations have in common the following inequalities:

$$\sum_{i \in L} x_{ij} = 1, \forall j \in X, \tag{1}$$

which means that every point j in X is assigned to a center location,

$$x_{ij} \leq y_i, \forall i \in L, j \in P, \tag{2}$$

i.e. if we assign j to i, then i must be open,

$$y_i, x_{ij} \in \{0, 1\}, \forall i \in L, j \in X, \tag{3}$$

the integrality constraints, and

$$\sum_{i \in L} y_i \leq k, \tag{4}$$

which limit the number of opened centers to k.

For the k-center case, the integer LP has no objective function. The idea is to guess the optimum value τ [BGK$^+$19]: Given τ, we construct a threshold graph $G_\tau = (X \cup L, E_\tau)$ on the sets of points and locations, where a connection between $i \in L$ and $j \in X$ is added iff the distance between i and j is within the threshold τ, i.e., $(i, j) \in E_\tau \Leftrightarrow d(i, j) \leq \tau$. Then, we need to make sure points are not assigned to centers outside their range, and this is achieved through

$$x_{ij} = 0, \forall i \in L, j \in P, (i, j) \notin E_\tau. \tag{5}$$

For the k-median case, we use the following objective function (6):

$$\min \sum_{i \in L, j \in X} x_{ij} d(i, j). \tag{6}$$

Replace (3) with $y_i, x_{ij} \in [0, 1], \forall i \in L, j \in X$ and we have the relaxed LP from the integer LP. To incorporate fairness, we add the constraints (7) with respect to the upper and lower bounds:

$$l_h \sum_{j \in X} x_{ij} \leq \sum_{col(p_j) = col_h} x_{ij} \leq u_h \sum_{j \in X} x_{ij}, \forall i \in L, col_h \in Col. \tag{7}$$

(1)-(5) and (7) represent the fair k-center problem, and (1)-(4) and (6)-(7) represent the fair k-median problem. We aim to find an optimal solution $\{x, y\}$ for both problems.

For a set X', we denote $mass_h(X') = |col_h(X')|$ as the mass of color col_h in X'. For a possibly fractional LP solution $\{x, y\}$, we extend this notion: $mass_h(x, i) = \sum_{j \in col_h(X)} x_{ij}$. We denote the total mass assigned to i in $\{x, y\}$ by $mass(x, i) = \sum_{j \in X} x_{ij}$. Now we can define essential fairness:

Definition 5 (Equivalent to Definition 6 of [BGK+19])**.** *Let I be an instance of a fair clustering problem and let $\{x, y\}$ be an integral but not necessarily fair solution to I. We say that $\{x, y\}$ is* essentially fair *if there exists a fractional fair solution $\{x', y'\}$ for I such that $\forall i \in L$:*

$$\lfloor mass_h(x', i) \rfloor \leq mass_h(x, i) \leq \lceil mass_h(x', i) \rceil, \forall col_h \in Col$$

and

$$\lfloor mass(x', i) \rfloor \leq mass(x, i) \leq \lceil mass(x', i) \rceil.$$

For essentially fair clustering, we employ approximation algorithms for (unfair) clustering problems as a black-box and transform their output into essentially fair solutions. We are going to assume that we are given two solutions, an integral unfair solution, and a fractional fair solution. The first step is to find a fractional fair assignment to the centers of the integral solution with a reasonable cost.

5.2 Combining two solutions

Let $\{x^{LP}, y^{LP}\}$ be an optimal solution to the LP with the property that all assignments are fair but the centers could be fractionally open and the points could be fractionally assigned to multiple centers, and let c_f be the value of this solution. Let $\{\bar{x}, \bar{y}\}$ be an integral solution to the LP that may violate the fairness constraints (7), and let c_i be the value of this solution. We denote the cost of the optimal integral solution to the LP by c. Our goal now is to combine $\{x^{LP}, y^{LP}\}$ and $\{\bar{x}, \bar{y}\}$ into a third solution $\{\hat{x}, \hat{y}\}$ such that the cost of $\{\hat{x}, \hat{y}\}$ is bounded by $O(c_f + c_i) \subseteq O(c)$ [BGK$^+$19]. Furthermore, the entries of $\hat{y}$ shall be integral.

Let S be the set of opened centers in $\{\bar{x}, \bar{y}\}$. For any $j \in X$, denote the center in S closest to j by $\bar{\phi}(j)$ (different notation from that in Section 2). Keep in mind that L may or not be a subset of X depending on the specific problem we are dealing with. We extend the $\bar{\phi}(j)$ function as follows. Let $i \in L \setminus X$ be any center, and let $\hat{j}$ be the closest point to i in X. Then we set $\bar{\phi}(i)$ to $\bar{\phi}(\hat{j})$. This means i is assigned to the center in S closest to the point in X that is closest to i. Finally, for any $i \in S$, let $\bar{C}(i) = \bar{\phi}^{-1}(i)$ be the set of all points and centers assigned to i by $\bar{\phi}$.

The following lemma effectively completes the step of combining the integral unfair solution with the fractional fair solution.

Lemma 13 (Equivalent to Lemma 7 of [BGK$^+$19])**.** *Let* $\{x^{LP}, y^{LP}\}$ *and* $\{\bar{x}, \bar{y}\}$ *be* *two solutions to the LP, where $\{\bar{x}, \bar{y}\}$ is integral but may violate inequality (7). Then the solution defined by*

$$\begin{aligned} \hat{y} &= \bar{y}, & & \\ \hat{x}_{ij} &= \sum\nolimits_{i' \in \bar{C}(i)} x_{i'j}^{LP}, & & \forall i \in S, j \in X, \\ \hat{x}_{ij} &= 0, & & \forall i \notin S, j \in X \end{aligned}$$

is fair, $\hat{y}$ is integral, and the cost $\hat{c}$ of $\{\hat{x}, \hat{y}\}$ is bounded by $c_f + c_i$ for k-center and by $2c_f + c_i$ for k-median.

Proof sketch. The definition of $\{\hat{x}, \hat{y}\}$ means that for every (fractional) assignment from a point j to a center i', we can shift this assignment to i, so from the perspective of i we collect all fractional assignments to centers in $\bar{C}(i)$ and consolidate them at i.

$\hat{y}$ is integral since $\bar{y}$ is integral. Next, we observe that $\{\hat{x}, \hat{y}\}$ satisfies fairness constraints because $\{x^{LP}, y^{LP}\}$ respects inequality (7), and we are moving all assignments from a center i' to the same center $\bar{\phi}(i')$. This shifting operation preserves fairness. The costs can be verified through straightforward (β-relaxed) triangular inequality arguments [BGK+19]. $\square$

5.3 Rounding the x-variables

The x-variables after the first step are not necessarily integral, so we need to round them here. Let $j \in X$ be a point that is fractionally assigned to a set of centers $L' \subseteq L$.

First we observe that for the k-center problem, we can shift mass from an assignment of j to $i' \in L'$ to an assignment of j to $i'' \in L'$ while preserving the objective. We say that such objectives are *reassignable* [BGK+19] (k-supplier also has this property).

For the k-median problem, the distances influence the cost in the form $\sum_{i\in L, j\in X} c_{ij}\cdot x_{ij}$ for some positive real number c_{ij}. We say that such objectives are *separable* in that the distances are a separate part of the total cost [BGK+19] (Facility location and k-means also fall into this category.)

Lemma 14 (Equivalent to Lemma 8 of [BGK+19])**.** *Let $\{x, y\}$ be an α-approximate fractional solution for a fair clustering problem such that all y_i are integral. Then we can obtain an α-approximate integral solution $\{x', y'\}$ with an additive fairness violation of at most one in time $O(poly(|S| + |X|))$.*

Proof sketch. We create our rounded solution $\{x', y'\}$ through min-cost flow (MCF) operations. We define a min-cost flow instance $(G = (V, A), c, b)$ with unit capacities and costs c on the edges as well as balances b on the nodes. We begin by defining a graph $G_h = (V_h, A_h)$ for every color index $h \in [g]$ with

$$V^h = V_S^h \cup V_X^h, \quad V_S^h = \{v_i^h : i \in S\}, \quad V_X^h = \{v_j^h : j \in col_h(X)\},$$

$$A^h = \{(v_j^h, v_i^h) : i \in S, j \in col_h(X), x_{ij} > 0\},$$

and costs c^h defined by $c_a^h = c_{ij}$ for $a = (v_j^h, v_i^h) \in A^h, i \in S, j \in col_h(X)$ and balances b by

$$b_v^h = \begin{cases} 1 & v \in V_X^h \\ -\lfloor mass_h(x, i)\rfloor & v = v_i^h \in V_S^h \end{cases}.$$

We then use G_h to define $G = (V, A)$:

$$V = \{t\} \cup V_S \cup \bigcup_{h\in Col} V^h, \quad V_S = \{v_i : i \in S\},$$

$$A = \bigcup_{h \in Col} A^h \cup \{(v_i^h, v_i) : i \in S, h \in Col, mass_h(x,i) - \lfloor mass_h(x,i) \rfloor > 0\}$$

$$\cup \{(v_i, t) : i \in S, mass(x,i) - \lfloor mass(x,i) \rfloor > 0\},$$

as well as costs c defined by $c_a = c_a^h$ for $a \in A^h$ and 0 otherwise. Balances b are defined by

$$b_v = \begin{cases} b_v^h & v \in V^h, h \in Col \\ -B_i & v = v_i \in V_S \end{cases}$$

and $b_t = -B$, where $B_i = \lfloor mass(x,i) \rfloor - \sum_{h \in Col} \lfloor mass_h(x,i) \rfloor$ and $B = |X| - \sum_{i \in S} \lfloor mass(x,i) \rfloor$.

See Figure 3 for an example G used to round the x-variables.

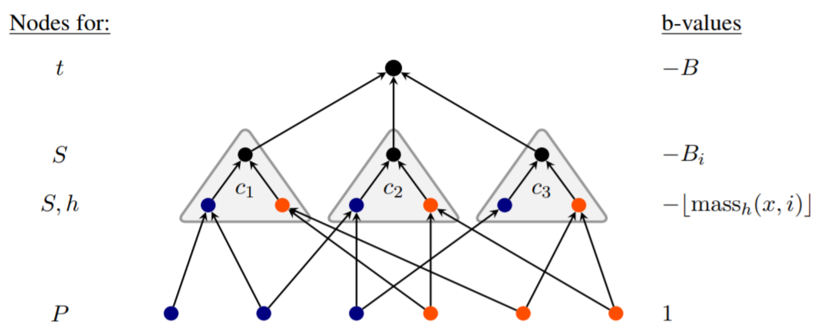

Figure 3: Example for the graph G used in rounding the x-variables ($P = X$).

We make the following observations about separable objectives (k-median) [BGK+19]:

1. All capacities, costs and balances and B and B_i for all $i \in S$ are all integers. Consequently, there are integral optimal solutions for the MCF instance (G, c, b).

2. A feasible solution for (G, c, b) is found in $\{x, y\}$ by defining a flow x in G in the

following way:

$$x_a = \begin{cases} x_{ij} & a = (v_j^h, v_j^h) \in A^h, j \in P, i \in S \\ mass_h(x,i) - \lfloor mass_h(x,i) \rfloor & a = (v_i^h, v_i) \in A, h \in Col, i \in S \\ mass(x,i) - \lfloor mass(x,i) \rfloor & a = (v_i, t) \in A, i \in S \end{cases}$$

Since $\{x, y\}$ is a fractional solution, x satisfies capacity and non-negativity constraints. We have flow conservation as well.

3. Integral solution x to the MCF instance (G, c, b) leads to an integral solution $(\bar{x}, y)$ to the original clustering problem by setting $\bar{x}_{ij} = x_a$ for $a = (v_j^h, v_i^h) \in A^h$ if $j \in col_h(X), i \in S$. This incurs the additive fairness violation of at most one, for every $i \in S$ is guaranteed to have at least $\lfloor mass_h(x,i) \rfloor$ points of color h and at least $\lfloor mass(x,i) \rfloor$ points in total assigned to it.

In the case of reassignable objectives (k-center), we basically employ the same strategy as before, but instead of a min-cost flow problem we solve the transshipment problem $(G = (V, A), b)$ with unit capacities on the edges and balances b on the nodes [BGK+19]. Notice that the three observations from the previous case apply here as well, and reassignability guarantees that the cost does not increase. □

Lemma 13 and Lemma 14 combined lead to the following theorem.

Theorem 12 (Equivalent to Theorem 9 of [BGK+19])**.** *Black-box approximation for fair clustering gives essentially fair solutions with a cost of $c_f + c_i$ for k-center and $2c_f + c_i$ for k-median.*

We know that c_f is not more expensive than an optimal solution to the fair clustering problem. The best known approximation factor for k-center is 2 [Gon85] and for k-median $(1 + \sqrt{3} + \epsilon)$ [LS16]. We arrive at the following theorem.

Theorem 13. *Black-box approximation for fair clustering gives essentially fair solutions with an approximation factor of 3 for k-center and* $(3+\sqrt{3}+\epsilon)$ *for k-median.*

5.4 True approximation for k-center

We now extend our rounding technique for the k-center problem to the case of exact fairness, using a specific approximation algorithm to obtain true approximation for the fair clustering problem via rounding of the LP solution. That is, we consider the fair k-center problem with exact preservation of ratios and without any additive fairness violation.

We begin by choosing a set of centers. Rather than use an arbitrary algorithm for the standard k-center problem, we specifically look for nodes in the threshold graph $G_\tau = (X, E_\tau)$ (τ is a parameter of G_τ and the optimal value we seek) where $E_\tau = \{(i,j) : i \neq j \in X, d(i,j) \leq \tau\}$ that form a maximal independent set S in G_τ^2 [BGK+19]. (Here G_τ^t denotes the graph on X that connects all pairs of nodes with a distance at most t in G, and E_τ^t denotes the edge set of G_τ^t. We also assume in this subsection that G_τ is a connected graph.)

The procedure uses the approach by Khuller and Sussmann [KS00] to find S with the following property: There exists a rooted tree T spanning all the nodes in S such that two adjacent nodes in T are exactly distance 3 apart in G_τ. The procedure begins by choosing a root vertex $n \in X$ into S and marking every node within distance 2 of n (including itself). Until all the nodes in X are marked, it chooses an unmarked node u that is adjacent to a marked node v and marks all nodes within distance two of u. Notice that u is exactly at distance 3 from a node $u' \in S$ chosen earlier that led to v getting marked. The tree T over the nodes of S is thus defined implicitly.

Next we make the following observation [BGK+19]:

Observation 1. Let $m \in \mathbb{N}$ be the smallest integer such that for each color $col_h \in Col$ we have $r_h(X) = \frac{q_h}{m}$ for some $q_h \in \mathbb{N}$. Then for each cluster X_i in a fair clustering C of X with exact preservation of ratios, there exists a positive integer $i' \in \mathbb{N}_{\geq 1}$ such that X_i contains exactly $i' \cdot q_h$ points with color h and $i' \cdot m$ total points. Therefore every cluster must have at least q_h points of color col_h for each $col_h \in Col$.

We use **Observation 1** and the set of centers S to obtain the following adjusted LP formulation for the fractional fair k-center problem [BGK+19].

$$\sum_{i \in S} x_{ij} = 1, \qquad \forall j \in X \tag{8}$$

$$\sum_{j \in col_h(X)} x_{ij} = r_h(X) \sum_{j \in X} x_{ij}, \qquad \forall i \in S \tag{9}$$

$$\sum_{j \in col_h(X), (i,j) \in E_\tau^2} x_{ij} \geq q_h, \qquad \forall i \in S, h \in Col \tag{10}$$

$$x_{ij} = 0, \qquad \forall i \in S, j \in X, (i,j) \notin E_\tau^3 \tag{11}$$

$$0 \leq x_{ij} \leq 1, \qquad \forall i \in S, j \in X. \tag{12}$$

Here inequality (10) ensures that each cluster contains at least q_h points of color $h \in Col$.

Finally, the algorithm rounds a fractional solution for the above LP to an integral solution of cost at most 5τ in a procedure motivated by Cygan et al. in [CHK12]. Let $\alpha(i)$ denote the children of node $i \in S$ in the tree T. Define quantities $\Gamma(i)$ and

$\delta(i)$ as follows:

$$\Gamma(i) = \left\lfloor \frac{\sum_{j \in col_1(X)} x_{ij} + \sum_{i' \in \alpha(i)} \delta(i')}{q_1} \right\rfloor$$
$$\delta(i) = \sum_{j \in col_1(X)} x_{ij} + \sum_{i' \in \alpha(i)} \delta(i') - \Gamma(i).$$

For a leaf node i in T, $\alpha(i) = \emptyset$. Then $\Gamma(i)$ represents the number of color 1 points assigned to i rounded down to the nearest multiple of q_1 and $\delta(i)$ denotes the remainder. We want to reassign the remainder to the parent of i. For a non-leaf i', $\Gamma(i')$ denotes the number of color 1 points assigned to i' plus the remainder that all children of i' want to reassign to their parent i' rounded down to the nearest multiple of q_1 while $\delta(i')$ again denotes the remainder.

Think of the x_{ij} variables as encoding flow from a vertex j to a node $i \in S$. We call it a color-h flow if j has color h. We can re-route these flows (maintaining the ratio constraints) such that $\forall i \in S, j \in col_1(X), x_{ij}$ is equal to $\Gamma(i)$ which is an integral multiple of q_1.

Lemma 15 (Equivalent to Lemma 12 of [BGK$^+$19])**.** *There exists an integral assignment of all vertices with color 1 to centers in S in G_τ^5 which assigns $\Gamma(i)$ vertices with color 1 to each center $i \in S$.*

Proof sketch. Construct the following flow network: Form a bipartite graph using sets $col_1(X)$ and S, with an edge of capacity one between $j \in col_1(X)$ and $i \in S$ iff $(i, j) \in E_\tau^5$. Connect a source s with unit capacity edges to all vertices in $col_1(X)$ and each center $i \in S$ with capacity $\Gamma(i)$ to a sink t. We now show a feasible fractional flow of value $|col_1(X)|$ in this network.

For each leaf node i in T, assign $\Gamma(i)$ amount of color-1 flow from the total incoming color-1 flow $\sum_{j \in col_1(X)} x_{ij}$, from vertices that are at most distance three away from i in G_tau and propagate the remainder $\delta(i)$, which comes from vertices of distance two from i, upward to be assigned to the parent of i. This is always possible because of constraint (10). For every non-leaf node i, assign $\Gamma(i)$ amount of incoming color-1 flow from distance five vertices (including the color-1 flows propagated upward by its children) and propagate $\delta(i)$ amount of color-1 flow from distance two vertices (possible due to constraint (10)). Hence we have that every center has $\Gamma(i)$ amount of color-1 flow passing through it. $\square$

Lemma 16 (Equivalent to Lemma 13 of [BGK+19])**.** *For any assignment of a color-1 flow, there exists a reassignment of color-h flow between the same centers for all $h \in Col \setminus \{1\}$, such that the resulting fractional assignment of the vertices satisfies the fairness constraints at each center.*

Proof. Suppose f_1 amount of color-1 flow is reassigned from center i_0 to another center i_1. Reassign $f_h = r_h \cdot f_1/r_1$ amount of color-h flow from i_0 to i_1 for each color $h \in Col \setminus \{1\}$. This is possible due to constraint (9). It is easy to verify that the ratios at i_0 and i_1 remain unchanged, for the ratio of the reassigned flows is equal to the original ratio by construction. $\square$

From Lemma 15 and Lemma 16, there is a fair fractional assignment within distance 5τ such that all the color 1 assignments are integral and every center i has $\Gamma(i)$ color 1 vertices assigned to it. Since this assignment is fair, the total incoming color-h flow at each center can be written as $\Gamma(i)\frac{q_h}{q_1}$, which is integral for every center $i \in S$ and every color $h \in Col$.

Lemma 17 (Equivalent to Lemma 14 of [BGK+19])**.** *There exists an integral fair assignment in G_τ^5.*

From Lemma 15, Lemma 16 and Lemma 17, we have the main theorem of this subsection.

Theorem 14 (Equivalent to Theorem 15 of [BGK+19])). *There exists a 5-approximation for the fair k-center problem with exact preservation of ratios.*

Remark: This result may not appear to be an upgrade over Section 1, which offers a 4-approximation algorithm for the fair k-center problem. Keep in mind, however, that the 4-approximation algorithm does not guarantee exact ratio preservation like the 5-approximation algorithm in this section does. In this sense Theorem 14 does signal an improvement, with consideration of a more general variant of the fair k-center problem in the form of multiple colors and exact preservation of ratio.

6 Conclusion

Machine learning is a very active research area, with a long history of work in both theoretical and practical aspects. In the last decade, the potential harm of algorithmic bias against underrepresented groups has become a popular concern. The research community has responded with multiple approaches to both define and address such issues. There has been a growing interest in developing ”fair” solutions to problems in learning and optimization.

In this work we have focused on one narrow slice of this work: worst-case approximation bounds for clustering algorithms with added fairness constraints, where we define fairness in terms of protected populations that must be represented in each cluster. Even within this narrow scope, we have found multiple recent approaches and diverse results, as presented in the preceding sections.

There are several ways we could hope to extend on this work.

- First, there is a need for empirical studies in order to implement and benchmark the proposed methods, both to compare their performance on large data sets, and to evaluate any practical demand for such methods. It seems likely that the approximation performance of the algorithms presented here could be improved by some greedy local searching, and we want to confirm this through experiments.

- Second, we could look for more general algorithmic approaches (e.g. some variant of LP rounding) that might unify some of the disparate results presented here. New approaches might also allow other side constraints on the clustering problems, beyond the fairness constraints studied here.

- Third, we hope to come up with more general ways to incorporate fairness notions into the clustering framework. For example, instead of hard constraints, we could add an "unfairness" cost term to the classical objective, and try to minimize the resulting system by standard learning methods. This would allow the user to choose an acceptable trade-off between unfairness and other costs. Other notions of fairness in clustering may pay attention to issues of crowding, relative distance, or balancing the load placed on opened facilities.

- More immediately, we may also attempt to solve closely related problems using the proposed methods. For example, we are currently investigating the use of the fairlet method to solve fair variants of the facility location problem. Facility location resembles the clustering problems already discussed, except that instead of opening a fixed number k of facilities, each facility has an opening cost, and we are actively exploring the possibility of solving the fair facility location problem under the fairlet notion.

References

[AEKM20] Sara Ahmadian, Alessandro Epasto, Ravi Kumar, and Mohammad Mahdian. Fair correlation clustering. *arXiv:2002.02274*, 2020.

[ANFSW17] Sara Ahmadian, Ashkan Norouzi-Fard, Ola Svensson, and Justin Ward. Better guarantees for k-means and euclidean k-median by primal-dual algorithms. In Chris Umans, editor, *58th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2017)*, page 61–72. IEEE Computer Society, 2017. doi:10.1109/FOCS.2017.15.

[BBC04] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. In *Mach. Learn., 56(1-3)*, pages 89–113, 2004.

[BCFN19] Suman K. Bera, Deeparnab Chakrabarty, Nicolas J. Flores, and Maryam Negahbani. Fair algorithms for clustering. In *NeurIPS*, page 4955–4966, 2019.

[BGK+19] Ioana O. Bercea, Martin Groß, Samir Khuller, Aounon Kumar, Clemens Rösner, Daniel R. Schmidt, and Melanie Schmidt. On the cost of essentially fair clusterings. In *APPROX-RANDOM*, page 18:1–18:22, 2019.

[CHK12] Marek Cygan, MohammadTaghi Hajiaghayi, and Samir Khuller. LP rounding for k-centers with non-uniform hard capacities. In Venkatesan Guruswami, editor, *53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2012)*, page 273–282. IEEE Computer Society, 2012. doi:10.1109/FOCS.2012.63.

[CKLV17] Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. *arXiv:1802.05733*, 2017.

[CMSY15] Shuchi Chawla, Konstantin Makarychev, Tselil Schramm, and Grigory Yaroslavtsev. Near optimal LP rounding algorithm for correlation clustering on complete and complete k-partite graphs. In *STOC*, pages 291–228, 2015.

[DEFI06] Erik D. Demaine, Dotan Emanuel, Amos Fiat, and Nicole Immorlica. Correlation clustering in general weighted graphs. *Theoretical Computer Science*, 361(2-3):172–187, 2006.

[DHP+12] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard S. Zemel. Fairness through awareness. In *ITCS*, pages 214–226, 2012.

[FFM+15] Michael Feldman, Sorelle A. Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. Certifying and removing disparate impact. In *KDD*, pages 259–268, 2015.

[Gon85] Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985. doi:`10.1016/0304-3975(85)90224-5`.

[HS85] Dorit S. Hochbaum and David B. Shmoys. A best possible heuristic for the k-center problem. *Mathematics of Operations Research*, 10:180–184, 1985. doi:`10.1287/moor.10.2.180`.

[KLS12] Tamás Király, Lap Chi Lau, and Mohit Singh. Degree bounded matroids and submodular flows. *Combinatorica*, 32(6):703–720, 2012.

[KS00] Samir Khuller and Yoram J. Sussmann. The capacitated k-center problem. *SIAM Journal on Discrete Mathematics*, 13:403–418, 2000. doi:`10.1137/S0895480197329776`.

[LS16] Shi Li and Ola Svensson. Approximating k-median via pseudo-approximation. *SIAM Journal on Computing*, 45:530–547, 2016. doi:10.1137/130938645.

[Vaz03] Vijay V. Vazirani. *Approximation Algorithms*. Springer-Verlag Berlin Heidelberg GmbH, 2003.