On the cost of essentially fair clusterings

Ioana O. Bercea[1], Martin Groß[2], Samir Khuller[1], Aounon Kumar[1], Clemens Rösner[3], Daniel R. Schmidt[4], and Melanie Schmidt[3]

[1] Department of Computer Science, University of Maryland, College Park, USA
[2] RWTH Aachen, Germany
[3] University of Bonn, Germany
[4] University of Cologne, Germany

Abstract. Clustering is a fundamental tool used in data mining to partition data into groups or classes. It can be used to make decisions that apply to each point based on the group it belongs to. However, especially if the data is already biased, some protected (minority) group can be harmed in this process because the clustering algorithm does not give it adequate representation in desirable clusters.

At NIPS 2017, Chierichetti et. al. [14] proposed a model for *fair clustering* which requires the representation in each cluster to (approximately) preserve the fraction of each protected class. For the special case that there are only two protected classes, they developed a 4-approximation for the fair k-center problem and a $\mathcal{O}(t)$-approximation for the fair k-median problem, where t is a parameter for the fairness model. For multiple protected classes, the best known result so far is a 14-approximation for fair k-center [32].

We extend and improve the known results. Firstly, we give a 6-approximation for the fair k-center problem with multiple protected classes. Secondly, we propose a relaxed fairness notion under which we can give bicriteria constant-factor approximations for all of the classical clustering objectives k-center, k-supplier, k-median, k-means and facility location. All approximation algorithms are achieved by a single framework which takes an existing unfair solution and a fair LP solution and combines them into an essentially fair clustering by using a weakly supervised rounding scheme. This has the advantage that a fair clustering can be established belatedly, in a situation where the centers are already fixed.

1 Introduction

Imagine being in charge of a big effort to reorganizing school assignments in your city. Given a long list of children that will start school next year and a short list of all available teachers, the goal is to assign the students-to-be to (public) schools such that the maximum distance to the school is small. The school capacity is given by the number of its teachers: For each teacher, s students can be admitted.

This challenge can be solved by using an algorithm for the capacitated (metric) k-center problem. But after a first assignment attempt, a school calls and says: This assignment means that our new classes will have much more boys

than girls, and the school next door has much more girls than boys. We would prefer an assignment where the classes are more balanced. So a new challenge arises: Assign the children such that the ratio is (approximately) 1:1 between boys and girls, and minimize the maximum distance under this condition.[5] This can be modeled by the following combinatorial optimization problem: Say we are given a point set where each point has a binary attribute, e.g., say each point is either red or blue. Assume that half of the points are red. Compute a clustering where each cluster has an equal number of red and blue points, and minimize the maximum radius.

In this form, our example is a special case of the *fair k-center* problem, proposed by Chierichetti et. al. [14] in the context of maintaining fairness in unsupervised machine learning tasks. Their model is based on the concept of *disparate impact* (and the p%-rule). The input points are assumed to have a binary sensitive attribute modeled by two colors, and discrimination based on this attribute is to be avoided. Since preserving exact balance in each cluster may be very costly or even be impossible[6], the idea is to ensure that at least $1/t$ of the points of each cluster are of the minority color, where t is a parameter. A cluster with this property is called *fair*, and the fairness constraint can now be added to any clustering problem, giving rise to fair k-center, fair k-median, and so on. Chierichetti at. al. develop a 4-approximation for fair k-center and a $(t + 1 + \sqrt{3} + \epsilon)$-approximation for fair k-median.

It is easy to imagine that the fair clustering model can also be used to incorporate other aspects into our school assignment example aside from an even ratio of boys and girls, for example to mitigate effects of gentrification or segregation. For this purpose, a model with more than two attribute values is useful. Then the idea is that the ratio of every value should be preserved. More precisely, the ratio between the number of points with one specific value and the total number of points in the cluster must be in some given range. For example, if the allowed range is $[0.20, 0.25]$ for red points, we require that in each cluster, at least a fifth and at most a fourth of the points are red. This models well established notions of fairness (statistical parity, group fairness), which require that each cluster exhibits the same compositional makeup as the overall data with respect to a given attribute.

> Text about fairness models in machine learning, references to some survey.

Finding fair clusterings turns out to be a quite interesting challenge from the point of view of combinatorial optimization. As other clustering problems with side constraints, it loses the property that points can be assigned locally. But while many other constraint problems at least allow polynomial algorithms that assign points to given centers optimally, we show that even this restricted problem is NP-hard in the case of fair k-center.

Chierichetti et. al. tackle fair clustering problems by a two-step procedure: First, they compute a micro clustering into so-called *fairlets*, which are groups of points that are fair and can not further be split into true subsets that are also fair.

---

[5] Or, incorporating the capacities, ensure that the teacher:boys:girls ratio is $1{:}\frac{s}{2}{:}\frac{s}{2}$.
[6] Imagine a point set with 49 red and 51 blue points: This can not at all be divided into true subsets with exact the same ratio.

Secondly, representative points of the fairlets are clustered by an approximation algorithm for the unconstrained problem. Consider the special case of a point set with 1:1 ratio of red and blue points. Then a fairlet is a pair of one red and one blue point, and a good micro clustering can be found by computing a suitable bipartite matching between the two color classes.

The problem of computing good fairlets gets increasingly difficult when considering more general variants of the problem. For multiple colors and the special case of exact ratio preservation (i.e., for all colors, the allowed range for its ratio is one specific number), the fairlet computation problem can be reduced to a capacitated clustering problem (with some additional complications). This is in [32] to obtain a 14/15-approximation for fair k-center/k-supplier with multiple colors and exact ratio preservation.

In the full version of this paper, we give an extensive overview on the existing results and develop the fairlet approach further in order to explore its applicability for different variants of fair clustering. Two major issues arise: Firstly, capacitated clustering is not solved for all clustering objectives; indeed, finding a constant-factor approximation for k-median is a long-standing open problem. Thus, the fairlet approach does not help for fair k-median or k-means with multiple colors. Secondly, it is even unclear for k-center how to compute fairlets when there are more than two colors *and* we want to allow a range for the ratios of the different colors. In this situation, it is not even clear how a fairlet looks like, since fairlets of very different size and composition may satisfy the desired ratio.

The main contribution of this paper is a completely different approach. We start by computing a solution to the unconstrained problem *first*. Since this can be any solution, we say our algorithm is a *black-box* approximation. Based on the given solution, we derive a fair clustering solution with the same centers. This step is achieved by a technique that we call *weakly supervised LP rounding*: We solve an LP for the fair clustering problem and then combine it with the integral unfair solution by careful rounding. We use this method to prove the following statements.

Theorem 1. *There exists a 5/7-approximation for the fair k-center/k-supplier problem with exact preservation of ratios.*

Theorem 2. *Given a (potentially unfair) clustering $\mathcal{C}$ with centers S and an assignment ϕ of the input set P, we can compute a clustering $\mathcal{C}'$ with the same set of centers S and an assignment ϕ' which is essentially fair. The cost of $\mathcal{C}'$ is linear in the cost of $\mathcal{C}$ and the cost of an optimal fair clustering of P, for all objectives k-center, k-supplier, k-median, k-means and facility location.*

Corollary 1. *There exists a 3/5/3.488/4.675/62.856-approximation for the fair k-center/k-supplier/facility location/k-median/k-means problem.*

The term *essentially fair* refers to our notion of bicriteria approximation. It means the following. A cluster is *essentially fair* if by fractionally changing the amount of colors by a total amount of *at most* 1, we can get all ratios into their

allowed ranges. So this is a small additive fairness violation. For large clusters, an additive violation of 1 will translate into a negligible multiplicative violation. Of course if a cluster is very small or a color has a very small ratio, then a violation of 1 is large in multiplicative terms. However, consider an example with one majority color having 500 points and 50 colors having 10 point each. Then our model means that in a cluster of 100 points, at most one of the minority colors can vanish completely because the *overall* amount of violation is at most one point. So in order to be fair, the majority color can still at most have 51 points, and at least 49 of the minority colors are preserved.

Our results have two advantages. Firstly, we get results for a wide range of clustering problems, and these results improve previous results. For example, we get a 5-approximation for the fair k-center problem with exact ratio preservation, where the best known guarantee was 14. For fair k-median, we get a 4.675-approximation with a small additive violation. While this result is of a bicriteria nature, note that the best known result was a non-constant approximation of $\Theta(t)$, which was only applicable to the case of two colors. The new result works for multiple colors, gives a constant factor approximation, and the fairness violation is bounded by an amount of one point per cluster. Notice that a similar result can *not* be achieved by using bicriteria approximation algorithms for capacitated clustering. The reduction from capacitated clustering only works when the capacities are not violated. Furthermore, our algorithms also work in the setting with multiple colors and approximate ratio preservation, for which no previous approximation algorithms were known.

Secondly, the black-box approach has the advantage that fairness can be established belatedly, in a situation where the centers are already given. [17,34]. Consider our school example and notice that the schools themselves cannot be chosen. Our result says that if we allow a one-point violation of fairness per cluster, we can achieve an essentially fair clustering which is not much more expensive than a fair clustering where the centers were chosen with the fairness constraint at hand. So keeping the centers is not a big disadvantage! When converting a solution into an essentially fair one, we can also track how previous clusters change in order to become fair, and see how our choice of ranges for the fairness influences the *cost of fairness* which is incurred.

Additional related work The unconstrained k-center/k-supplier problem can be 2/3-approximated [18,20], and this is tight [21]. Facility location can be 1.488-approximated [28], which is very close to the known APX-hardness of 1.463 for the problem [19]. For k-median, a recent breakthrough has led to a 2.765-approximation [31,11], while the best hardness result lies below two [22]. The gap between best upper and lower bound is even larger for k-means, where a 6.357-approximation is the best known [4], and the newest hardness result is marginally above 1 [8,25]. The wide applicability of k-center has given rise to numerous variants that strive to incorporate useful constraints into the solution, such as capacity constraints [10,15,23], lower bounds on the size of each cluster [3,6] or allowing for outliers [12,16].

The facility location problem has also been studied under capacity constraints [2,7,9], with uniform lower bounds [5,33], and with outliers [12]. Much less is known for k-median and k-means. True constant-factor approximations so far exist only for the outlier constraint [13,24], and are not known for the variants with capacities or lower bounds. A major problem for obtaining constant factor approximations is that the natural LP for the problem has an unbounded integrality gap, which as we will see is also true for the LP with fairness constraints. Bicriteria approximations are known that either violate the capacity constraints [27,29,30] or the cardinality constraint

A clustering problem where the points have a color was considered by Li, Yi and Zhang [26]. They provided a 2-approximation for a constraint called *diversity*, which allows at most one point per color in each cluster. Fair clustering as considered in this paper is studied in [14] and [32].

Preliminaries

Points and locations. In all the problems considered here, we are given a set of n points P and a set of potential locations L. We allow L to be infinite (for instance, $L = \mathbb{R}^d$). The task is to open a subset $S \subseteq L$ of the locations and to assign each point in P to an open location.

Assigning a point to a location incurs a cost governed by a semi-metric[7] $d : (P \cup L) \times (P \cup L) \to \mathbb{R}_{\geq 0}$ on $P \cup L$ that fulfills an β-*relaxed triangle inequality*

$$d(x,z) \leq \beta(d(x,y) + d(y,z)) \quad \text{for all } x,y,z \in P \cup L \tag{1}$$

for some $\beta \geq 1$. We extend d to sets by defining $d(x,Y) := \min_{y \in Y} d(x,y)$ for $Y \subseteq P \cup L$, and $d(X,Y) := \min_{x \in X} d(x,Y)$ for $X, Y \subseteq P \cup L$. Furthermore, we have opening costs $f_i \geq 0$ for every potential location $c_i \in L$ or may have to guarantee that at most $k \in \mathbb{N}$ locations are opened. We assign points in P to locations in L via a mapping $\phi : P \to S$ that we have to compute together with choosing S. If the assignment ϕ is clear from the context, we refer to the set of all points assigned to a location $c_i \in S$ by $P(i) := \phi^{-1}(i)$.

Colors and fairness. Furthermore, we are given a set of *colors* $Col := \{col_1, \ldots, col_g\}$, and a coloring $col : P \to Col$ that assigns a color to each point $j \in P$. For any set of points $P' \subseteq P$ and any color $col_h \in Col$ we define $col_h(P') = \{j \in P' \mid col(j) = col_h\}$ to be the set of points colored with col_h in P' using a slight abuse of notation. We also call $\text{mass}_h(P') = |col_h(P')|$ the *mass* of color col_h in P' and we call $r_h(P') := \frac{|col_h(P')|}{|P'|}$ the *ratio* of col_h in P'. If an implicit assignment ϕ is clear from the context, we write $col_h(i)$ to denote the set of all points of a color $col_h \in Col$ assigned to an $i \in S$, i.e., $col_h(i) = col_h(P(i))$.

In our fair clustering problems, we want to preserve the ratios of colors found in P in our clusters. We distinguish two cases: *exact* preservation of ratios, and

---

[7] A semi-metric $d : X \times X \to \mathbb{R}_{\geq 0}$ fulfills $d(x,y) \geq 0$, $d(x,y) = 0 \Leftrightarrow x = y$ and $d(x,y) = d(y,x)$ for all $x,y \in X$.

relaxed preservation of ratios. For the exact preservation of ratios, we ask that all the points assigned to a location $i \in S$ have exactly the same ratios as P, i.e.,

$$r_h(P(i)) = r_h(P) \quad \text{for all } col_h \in Col \text{ and } i \in S. \tag{2}$$

For the relaxed preservation of ratios, we have upper and lower bounds on the ratio of a color in each cluster. More precisely, we are given $p_1^h, q_1^h, p_2^h, q_2^h \in \mathbb{N}$ for every $c_h \in Col$, and the ratios have to fulfill

$$\ell_h := \frac{p_1^h}{q_1^h} \le r_h(P(i)) \le \frac{p_2^h}{q_2^h} =: u_h \quad \text{for all } col_h \in Col \text{ and } i \in S. \tag{3}$$

The exact case is a special case of the relaxed case where we set $\ell_h = u_h = r_h(P)$ for every color $col_h \in Col$.

A set of points $P' \subseteq P$ is *exactly fair* if P' has the same ratio for every color as P, i.e., for each $col_h \in Col$ we have $r_h(P') = r_h(P)$. We say that P' is ℓ, u-*fair* or just *fair* for some $\ell = (\ell_1 = p_1^1/q_1^1, \ldots, \ell_g = p_1^g/q_1^g)$ and $u = (u_1 = p_2^1/q_2^1, \ldots, u_g = p_2^g/q_2^g)$ if we have $r_h(P') \in [\ell_h, u_h]$ for every color $col_h \in Col$. We say that a clustering (S, ϕ) is *essentially fair*, if there exists a fair fractional solution (x, y) such that (S, ϕ) can be obtained from (x, y) with an additive fairness violation of at most one (see Definition 2).

Objectives. We consider fair versions of several classical clustering problems, in particular, we consider fair versions of k-center, k-supplier, k-median, k-means and facility location. An instance I of one of these problems is given by $I := (P, L, col, d, f, k, \ell, u)$, and our goal is to choose a solution (S, ϕ) according to a problem-specific objective that we will discuss now.

- k-**center** and k-**supplier**: minimize the maximum distance between a point and its assigned location: $\min \max_{j \in P} d(j, \phi(j))$. In these problems, we have $f \equiv 0$ and d is a metric. Furthermore, in k-center, $L = P$, whereas in k-supplier , $L \ne P$ is some finite set.
- k-**median**: minimize $\sum_{j \in P} d(j, \phi(j))$, d is a metric, $f \equiv 0$ and $L \subseteq P$.
- k-**means**: minimize $\sum_{j \in P} d(j, \phi(j))$, where $P \subseteq \mathbb{R}^m$ for some $m \in \mathbb{N}$, $L = \mathbb{R}^m$ and $d(x, y) = ||y - x||^2$ is a semi-metric for $\beta = 2$ and $f \equiv 0$.
- **facility location**: minimize $\sum_{j \in P} d(j, \phi(j)) + \sum_{i \in S} f_i$, where $k = n$, d is a metric and L is a finite set.

The fair assignment problem. For all the objectives above, we call the subproblem of computing a cost-minimal fair assignment of points to given centers the *fair assignment problem*. In the full version of this paper, we show that it is NP-hard to approximate the fair assignment problem for k-center better than 3.

(I)LP formulations for fair clustering problems

Let $I = (P, L, col, d, f, k, \ell, u)$ be a problem instance for a fair clustering problem. We introduce a binary variable $y_i \in \{0, 1\}$ for all $i \in L$ that decides if i is opened,

i.e. $y_i = 1 \Leftrightarrow i \in S$. Similarly, we introduce binary variables $x_{ij} \in \{0,1\}$ for all $i \in L, j \in P$ with $x_{ij} = 1$ if j is assigned to i, i.e. $\phi(j) = i$. The following constraints are the base of the ILP formulations.

$$\sum_{i \in L} x_{ij} = 1 \; \forall j \in P \quad (4) \qquad x_{ij} \leq y_i \; \forall i \in L, j \in P \quad (5) \qquad y_i, x_{ij} \in \{0,1\} \; \forall i \in L, j \in P \quad (6)$$

Inequalities (4) make sure that each point j is assigned to a center while inequalities (5) ensure that if we assign j to some center i, then i must be open. We may restrict the number of open centers to k with

$$\sum_{i \in L} y_i \leq k \qquad (7)$$

For k-center and k-supplier, the objective is commonly encoded in the constraints of the problem, and the (I)LP has no objective function. The idea is to guess the optimum value τ. Since there is only a polynomial number of choices for τ, this is easily done. Given τ, we construct a *threshold graph* $G_\tau = (P \cup L, E_\tau)$ on the points and locations, where a connection between $i \in L$ and $j \in P$ is added iff i and j are close, i.e., $\{i,j\} \in E_\tau \Leftrightarrow d(i,j) \leq \tau$. Then, the we ensure that points are not assigned to centers outside their range:

$$x_{ij} = 0 \qquad \qquad \text{for all } i \in L, j \in P, \{i,j\} \notin E_\tau \qquad (8)$$

For the remaining clustering problems, we pick the adequate objective function from the following three:

$$\min \sum_{i \in L, j \in P} x_{ij} d(i,j) \quad (9) \qquad \min \sum_{i \in L, j \in P} x_{ij} d^2(i,j) \quad (10) \qquad \min \sum_{i \in L, j \in P} x_{ij} d(i,j) + \sum_{i \in L} y_i f_i \quad (11)$$

We now have all necessary constraints and objectives. For k-center and k-supplier, we use inequalities (4)-(8), no objective, and define the optimum to be the smallest τ for which the ILP has a solution. We get k-median and k-means by combining inequalities (4)-(7) with (9) and (10), respectively, and we get facility location by combining (4)-(6) with the objective (11). LP relaxations arise from all ILP formulations by replacing (6) by $y_i, x_{ij} \in [0,1]$ for all $i \in L, j \in P$. To create the fair variants of the ILP formulations, we add fairness constraints modeling the upper and lower bound on the balances.

$$\ell_h \sum_{j \in P} x_{ij} \leq \sum_{\substack{col(p_j)=col_h}} x_{ij} \; \forall i \in L, h \in \{1,\ldots,g\} \quad (12) \qquad u_h \sum_{j \in P} x_{ij} \geq \sum_{\substack{col(p_j)=col_h}} x_{ij} \; \forall i \in L, h \in \{1,\ldots,g\} \quad (13)$$

Although very similar to the canonical clustering LPs, the resulting LPs become much harder to round even for k-center with two colors. We show that the integrality gap is $\Omega(n)$ in Section C. This is also true for the formulation considered by Chierichetti et al [14].

Essential fairness. We first extend natural definitions to fractional solutions.

Definition 1 (Mass and ratio of a color assigned to a center). *Let (x, y) be a fractional solution to an instance of a clustering problem. We define the* mass *of a color $col_h \in Col$ assigned to a center $i \in S$ to be $\mathrm{mass}_h(x, i) := \sum_{j \in col_h(P)} x_{ij}$ and the* mass *assigned to a center $i \in S$ to be $\mathrm{mass}(x, i) = \sum_{j \in P} x_{ij}$.*

With this notation, we can now formalize our notion of *essential fairness*.

Definition 2 (Essential fairness). *Let I be an instance of a fair clustering problem, let (x, y) be a fractional solution for I and $(x', y') := \vartheta(x, y)$ be an integral solution to I. We say that ϑ is* essentially fair *if $\forall i \in L$ with $y_i > 0$:*

$$\lfloor \mathrm{mass}_h(x, i) \rfloor \leq \mathrm{mass}_h(x', i) \leq \lceil \mathrm{mass}_h(x, i) \rceil \qquad \forall col_h \in Col \quad (14)$$

$$and \quad \lfloor \mathrm{mass}(x, i) \rfloor \leq \mathrm{mass}(x', i) \leq \lceil \mathrm{mass}(x, i) \rceil. \qquad (15)$$

2   Essential fair clusterings via black-box approximation

For essentially fair clustering, we give a powerful framework that employs approximation algorithms for (unfair) clustering problems as a black-box and transforms their output into an essentially fair solution. In this framework, we start by computing an approximate solution for the standard variant of the clustering problem at hand. Next, we solve the LP for the fair variant of the clustering problem. Now we have an integral unfair solution, and a fractional fair solution. Our final and most important step is to combine these two solutions into an integral and essentially fair solution. It consists of two conceptual sub-steps: Firstly, we show that it is possible to find a fractional fair assignment to the centers of the integral solution that is sufficiently cheap. Secondly, we round the assignment. This last sub-step introduces the potential fairness violation of one point per color per cluster. Algorithms 1 and 2 summarize this framework.

We show that this approach yields constant-factor approximations with fairness violation for all mentioned clustering objectives. The description will be neutral whenever the objective does not matter. Thus, descriptions like *the LP* mean the appropriate LP for the desired clustering problem. When the problem gets relevant, we will specifically discuss the distinctions. Notice that for all clustering problems defined in Section 1, P and L are finite except for k-means. However, for the k-means problem, we can assume that $L = P$ if we accept an additional factor of 2 in the approximation guarantee (Lemma 12 in the appendix). Thus, we assume in the following that L and P are finite sets. Indeed, we even assume at least $L \subseteq P$ for all problems except k-supplier and facility location.

2.1   Step 1: Obtaining a fair solution with integral y

In the first step, we assume that we are given two solutions. Let (x^{LP}, y^{LP}) be an optimal solution to the LP. This solution has the property that the assignments

to all centers are fair, however, the centers may be fractionally open and the points may be fractionally assigned to several centers. Let c^{LP} be the objective value of this solution. For k-supplier and k-center, it is the smallest τ for which the LP is feasible, for the other objectives, it is the value of the LP. We denote the cost of the best *integral* solution to the LP by c^*. We know that $c^{LP} \leq c^*$.

Let $(\bar{x}, \bar{y})$ be any integral solution to the LP that may violate fairness, i.e., inequalities (12) and (13), and let $\bar{c}$ be the objective value of this solution. We think of $(\bar{x}, \bar{y})$ as being a solution of an α-approximation algorithm for the standard (unfair) clustering problem for some constant α. Since the unconstrained version can only have a lower optimum cost, we then have $\bar{c} \leq \alpha \cdot c^*$.

Our goal is now 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 $O(c^{LP} + \bar{c}) \subseteq O(c^*)$. Furthermore, the entries of $\hat{y}$ shall be integral. The entries of $\hat{x}$ may still be fractional after step 1.

Let S be the set of centers that are open in $(\bar{x}, \bar{y})$. For all $j \in P$, we use $\bar{\phi}(j)$ to denote the center in S closest to j, i.e., $\bar{\phi}(j) = \arg\min_{i \in S} \mathrm{d}(j, i)$ (ties broken arbitrarily). Notice that the objective value of using S with assignment $\bar{\phi}$ for all points in P is at most $\bar{c}$, since assigning to the closest center is always optimal for the standard clustering problems without fairness constraint.

Depending on the objective, L is a subset of P or not, i.e., $\bar{\phi}$ is not necessarily defined for all locations in L. We then extend $\bar{\phi}$ in the following way. Let $i \in L \backslash P$ be any center, and let j^* be the closest point to it in P. Then we set $\bar{\phi}(i) := \bar{\phi}(j^*)$, i.e., i is assigned to the center in S which is closest to the point in P which is closest to i. Finally, let $\bar{C}(i) = \bar{\phi}^{-1}(i)$ be the set of all points and centers assigned to i by $\bar{\phi}$. We show the following lemma (proof is in the appendix).

Lemma 1. *Let (x^{LP}, y^{LP}) and $(\bar{x}, \bar{y})$ be two solutions to the LP, where $(\bar{x}, \bar{y})$ may violate inequalities (12) and (13), but is integral. Then the solution defined by $\hat{y} := \bar{y}$ and*

$$\hat{x}_{ij} := \sum_{i' \in \bar{C}(i)} x_{i'j}^{LP} \quad \text{for all } i \in S, j \in P, \qquad \hat{x}_{ij} := 0 \quad \text{for all } i \notin S, j \in P.$$

satisfies inequalities (12) and (13), $\hat{y}$ is integral, and the cost $\hat{c}$ of $(\hat{x}, \hat{y})$ is bounded by $c^{LP} + \bar{c}$ for k-center, by $2 \cdot c^{LP} + \bar{c}$ for k-supplier, k-median, and facility location, and by $12 \cdot c^{LP} + 8 \cdot \bar{c}$ for k-means.

2.2   Step 2: Rounding the x-variables

For rounding the x-variables, we need to distinguish between two cases of objectives. Let $j \in P$ be a point that is fractionally assigned to some centers $L_j \subseteq L$.

First, we have objectives where we can shift mass from an assignment of j to $i' \in L_j$ to an assignment of j to $i'' \in L_j$ without modifying the objective. We say that such objectives are *reassignable* (in the sense that we can reassign j to centers in L_j without changing the cost). k-center and k-supplier have this property.

Second, we have objectives where the assignment cost is separable, i.e., where the distances influence the cost via a term of the form $\sum_{i \in L, j \in P} c_{ij} \cdot x_{ij}$ for some

$c_{ij} \in \mathbb{R}_{\geq 0}$. We call such objectives *separable*. Facility location, k-median and k-means fall into the this category.

Lemma 2. *Let (x, y) be an α-approximative fractional solution for a fair clustering problem with the property that all $y_i, i \in L$ are integral. Then we can obtain an α-approximative integral solution (x', y') with an additive fairness violation of at most one in time $O(\text{poly}(|S| + |P|))$, with $S := \{i \in L \mid y_i \geq 1\}$ being the set of locations that are opened in (x, y).*

Proof. We create our rounded α-approximate integral solution (x', y') by min-cost flow computations. We begin by constructing a min-cost flow instance which depends on our starting solution (x, y) as well as on the objective of the problem we are studying.

We define a min-cost flow instance $(G = (V, A), c, b)$ (also see Figure 1) 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 $h \in Col$ with

$$V^h := V_S^h \cup V_P^h, \quad V_S^h := \left\{ v_i^h \mid i \in S \right\}, \quad V_P^h := \left\{ v_j^h \mid j \in col_h(P) \right\},$$
$$A^h := \left\{ (v_j^h, v_i^h) \mid i \in S, j \in col_h(P) : x_{ij} > 0 \right\},$$

as well as costs c^h by $c_a^h := c_{ij}$ for $a = (v_j^h, v_i^h) \in A^h, i \in S, j \in col_h(P)$ and balances b^h by $b_v^h := 1$ if $v \in V_P^h$ and $b_v^h := -\lfloor \text{mass}_h(x, i) \rfloor$ if $v = v_i^h \in V_S^h$. We use the graphs G_h to define $G = (V, A)$ by

$$V := \{t\} \cup V_S \cup \bigcup_{h \in Col} V^h, \quad V_S := \{v_i \mid i \in S\}$$
$$A := \bigcup_{h \in Col} A^h \cup \left\{ (v_i^h, v_i) \mid i \in S, h \in Col : \text{mass}_h(x, i) - \lfloor \text{mass}_h(x, i) \rfloor > 0 \right\}$$
$$\cup \left\{ (v_i, t) \mid i \in S : \text{mass}(x, i) - \lfloor \text{mass}(x, i) \rfloor > 0 \right\},$$

together with costs c of $c_a := c_a^h$ for $a \in A^h$ and 0 otherwise, and balances b of $b_v := b_v^h$ if $v \in V^h$ for some $h \in Col$, $b_v := -B_i$ if $v = v_i \in V_S$ and $b_t := -B$ with $B_i = \lfloor \text{mass}(x, i) \rfloor - \sum_{h \in Col} \lfloor \text{mass}_h(x, i) \rfloor$ and $B := |P| - \sum_{i \in S} \lfloor \text{mass}(x, i) \rfloor$.

Separable objectives – k-median and k-means. We make the following observations.

1. B and B_i are integers for all $i \in S$, and so are all capacities, costs and balances. Consequently, there are integral optimal solutions for the min-cost flow instance (G, c, b),
2. (x, y) induces a feasible solution for (G, c, b), by defining a flow x in G as follows:

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

Since (x, y) is a fractional solution, x satisfies capacity and non-negativity constraints because $x_{ij} \in [0, 1]$ for all $i \in L, j \in P$ and $\mathrm{mass}_h(x, i) - \lfloor \mathrm{mass}_h(x, i) \rfloor, \mathrm{mass}(x, i) - \lfloor \mathrm{mass}(x, i) \rfloor \in [0, 1]$ for all $i \in S$ and $col_h \in Col$ as well. We have flow conservation since the fractional solution needs to assign all points, and the flow of the edges (v_i^h, v_i) and (v_i, t) as well as the demand of v_i and t are chosen in such a way that we have flow conservation for all the other nodes as well.
3. Integral solutions x to the min-cost flow instance (G, c, b) induce 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 $j \in col_h(P), i \in S$. Since the flow x is integral, this gives us an integral assignment of all points to centers which have been opened, since y was already integral before this step.

   This incurs the additive fairness violation of at most one, since every $i \in S$ is guaranteed by our balances to have at least $\lfloor \mathrm{mass}_h(x, i) \rfloor$ points of color $h \in Col$ and at least $\lfloor \mathrm{mass}(x, i) \rfloor$ points in total assigned to it. Since there is at most one outgoing arc of unit capacity (v_i^h, v_i) and (v_i, t) for an $i \in S$ if $\mathrm{mass}_h(x, i) - \lfloor \mathrm{mass}_h(x, i) \rfloor > 0$, we have at most $\lceil \mathrm{mass}_h(x, i) \rceil$ points of color col_h and $\lceil \mathrm{mass}(x, i) \rceil$ total points assigned to i.

Together, this yields that computing a min-cost flow $\hat{x}$ for (G, c, b) followed by applying the third observation to $\hat{x}$ yields a solution $(\hat{x}, y)$ to the clustering with an additive fairness violation of at most one.

Since (x, y) was inducing the fractional solution x with $\mathrm{cost}(x) = \mathrm{cost}(x, y)$ to the min-cost flow instances, and $\mathrm{cost}(x) \geq \mathrm{cost}(\hat{x})$ by construction we have $\mathrm{cost}(\hat{x}, y) \leq \mathrm{cost}(x, y)$.

Reassignable objectives – k-center and k-supplier. In the case of reassignable objectives, we do not have to care about costs, as long as the reassignments happen to centers in L_j for all points $j \in P$. We essentially use 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. Notice that the three observations from the previous case apply here as well, and reassignability guarantees that the cost does not increase.

Lemmas 1 and 2 then lead directly to Theorem 2, formulated in more detail in the following theorem.

Theorem 3. *Black-box approximation for fair clustering gives essentially fair solutions with a cost of $c^{LP} + \bar{c}$ for k-center, $2c^{LP} + \bar{c}$ for k-supplier, k-median and facility location, and $12c^{LP} + 8\bar{c}$ for k-means where c^{LP} is the cost of an optimal solution to the fair LP relaxation and $\bar{c}$ is the cost of the unfair clustering solution.*

We know that c^{LP} is not more expensive than an optimal solution to the fair clustering problem. If we use an α-approximation to obtain the unfair clustering solution, we have that $\bar{c}$ is at most α times the cost of an optimal solution to the fair clustering problem. Currently, the best know approximation factors are 2 for k-center [18,20], 3 for k-supplier [20], 1.488 for facility location [28], 2.675 for k-median [11,31] and 6.357 for k-means [4], yielding the following theorem.

Theorem 4. *Black-box approximation for fair clustering gives essentially fair solutions with an approximation factor of 3 for k-center, 5 for k-supplier, 4.675 for k-median, 3.488 for facility location, and 62.856 for k-means.*

References

1. Karen Aardal, Pieter L. van den Berg, Dion Gijswijt, and Shanfei Li. Approximation algorithms for hard capacitated k-facility location problems. *European Journal of Operational Research*, 242:358–368, 2015. `doi:10.1016/j.ejor.2014.10.011`.
2. Ankit Aggarwal, Anand Louis, Manisha Bansal, Naveen Garg, Neelima Gupta, Shubham Gupta, and Surabhi Jain. A 3-approximation algorithm for the facility location problem with uniform capacities. *Mathematical Programming*, 141:527–547, 2013. `doi:10.1007/s10107-012-0565-4`.
3. Gagan Aggarwal, Rina Panigrahy, Tomás Feder, Dilys Thomas, Krishnaram Kenthapadi, Samir Khuller, and An Zhu. Achieving anonymity via clustering. *ACM Transactions on Algorithms*, 6:49:1–49:19, 2010. `doi:10.1145/1798596.1798602`.
4. 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, *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science (FOCS 2017)*, pages 61–72. IEEE Computer Society, 2017. `doi:10.1109/FOCS.2017.15`.
5. Sara Ahmadian and Chaitanya Swamy. Improved approximation guarantees for lower-bounded facility location. In Thomas Erlebach and Giuseppe Persiano, editors, *10th International Workshop on Approximation and Online Algorithms (WAOA 2012)*, volume 7846 of *Lecture Notes in Computer Science (LNCS)*, pages 257–271. Springer Berlin Heidelberg, 2012. `doi:10.1007/978-3-642-38016-7_21`.
6. Sara Ahmadian and Chaitanya Swamy. Approximation algorithms for clustering problems with lower bounds and outliers. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 69:1–69:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016. `doi:10.4230/LIPIcs.ICALP.2016.69`.
7. Hyung-Chan An, Mohit Singh, and Ola Svensson. LP-based algorithms for capacitated facility location. *SIAM Journal on Computing*, 46:272–306, 2017. `doi:10.1137/151002320`.
8. Pranjal Awasthi, Moses Charikar, Ravishankar Krishnaswamy, and Ali Kemal Sinop. The hardness of approximation of euclidean k-means. In Lars Arge and János Pach, editors, *31st International Symposium on Computational Geometry (SoCG 2015)*, volume 34 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 754–767. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. `doi:10.4230/LIPIcs.SOCG.2015.754`.
9. Manisha Bansal, Naveen Garg, and Neelima Gupta. A 5-approximation for capacitated facility location. In Leah Epstein and Paolo Ferragina, editors, *Algorithms – ESA 2012*, volume 7501 of *Lecture Notes in Computer Science (LNCS)*, pages 133–144. Springer Berlin Heidelberg, 2012. `doi:10.1007/978-3-642-33090-2_13`.
10. Judit Bar-Ilan, Guy Kortsarz, and David Peleg. How to allocate network centers. *Journal of Algorithms*, 15:385–415, 1993. `doi:10.1006/jagm.1993.1047`.

11. Jarosław Byrka, Thomas Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k-median and positive correlation in budgeted optimization. *ACM Transactions on Algorithms*, 13:23:1–23:31, 2017. `doi:10.1145/2981561`.

12. Moses Charikar, Samir Khuller, David M. Mount, and Giri Narasimhan. Algorithms for facility location problems with outliers. In S. Rao Kosaraju, editor, *Proceedings of the Twelfth Annual Symposium on Discrete Algorithms (SODA 2001)*, pages 642–651, 2001. URL: `http://dl.acm.org/citation.cfm?id=365411.365555`.

13. Ke Chen. A constant factor approximation algorithm for k-median clustering with outliers. In Shang-Hua Teng, editor, *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2008)*, pages 826–835. SIAM, 2008. URL: `http://dl.acm.org/citation.cfm?id=1347082.1347173`.

14. Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017 (NIPS 2017)*, pages 5036–5044, 2017. URL: `http://papers.nips.cc/paper/7088-fair-clustering-through-fairlets`.

15. 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)*, pages 273–282. IEEE Computer Society, 2012. `doi:10.1109/FOCS.2012.63`.

16. Marek Cygan and Tomasz Kociumaka. Constant factor approximation for capacitated k-center with outliers. In Ernst W. Mayr and Natacha Portier, editors, *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 251–262. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014. `doi:10.4230/LIPIcs.STACS.2014.251`.

17. Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard S. Zemel. Fairness through awareness. In Shafi Goldwasser, editor, *Innovations in Theoretical Computer Science 2012 (ITCS 2012)*, pages 214–226. ACM, 2012. `doi:10.1145/2090236.2090255`.

18. 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`.

19. Sudipto Guha and Samir Khuller. Greedy strikes back: Improved facility location algorithms. *Journal of Algorithms*, 31:228–248, 1999. `doi:10.1006/jagm.1998.0993`.

20. Dorit S. Hochbaum and David B. Shmoys. A unified approach to approximation algorithms for bottleneck problems. *Journal of the ACM*, 33(3):533–550, 1986. `doi:10.1145/5925.5933`.

21. Wen-Lian Hsu and George L. Nemhauser. Easy and hard bottleneck location problems. *Discrete Applied Mathematics*, 1:209–215, 1979. `doi:10.1016/0166-218X(79)90044-1`.

22. Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In John H. Reif, editor, *Proceedings on 34th Annual ACM Symposium on Theory of Computing (STOC 2002)*, pages 731–740. ACM, 2002. `doi:10.1145/509907.510012`.

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

24. Ravishankar Krishnaswamy, Shi Li, and Sai Sandeep. Constant approximation for k-median and k-means with outliers via iterative rounding. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2018)*, pages 646–659. ACM, 2018. `doi:10.1145/3188745.3188882`.

25. Euiwoong Lee, Melanie Schmidt, and John Wright. Improved and simplified inapproximability for k-means. *Information Processing Letters*, 120:40–43, 2017. `doi:10.1016/j.ipl.2016.11.009`.

26. Jian Li, Ke Yi, and Qin Zhang. Clustering with diversity. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *Automata, Languages and Programming (ICALP 2010)*, volume 6198 of *Lecture Notes in Computer Science (LNCS)*, pages 188–200. Springer Berlin Heidelberg, 2010. `doi:10.1007/978-3-642-14165-2_17`.

27. Shanfei Li. An improved approximation algorithm for the hard uniform capacitated k-median problem. In Klaus Jansen, José D. P. Rolim, Nikhil R. Devanur, and Cristopher Moore, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*, volume 28 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 325–338. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014. `doi:10.4230/LIPIcs.APPROX-RANDOM.2014.325`.

28. Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Information and Computation*, 222:45–58, 2013. `doi:10.1016/j.ic.2012.01.007`.

29. Shi Li. Approximating capacitated k-median with $(1 + \epsilon)k$ open facilities. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016)*, pages 786–796. SIAM, 2016. `doi:10.1137/1.9781611974331.ch56`.

30. Shi Li. On uniform capacitated k-median beyond the natural LP relaxation. *ACM Transactions on Algorithms*, 13:22:1–22:18, 2017. `doi:10.1145/2983633`.

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

32. Clemens Rösner and Melanie Schmidt. Privacy Preserving Clustering with Constraints. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 96:1–96:14. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018. `doi:10.4230/LIPIcs.ICALP.2018.96`.

33. Zoya Svitkina. Lower-bounded facility location. *ACM Transaction on Algorithms*, 6:69:1–69:16, 2010. `doi:10.1145/1824777.1824789`.

34. Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. Learning fair representations. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning (ICML 2013)*, volume 28 of *Proceedings of Machine Learning Research*, pages 325–333. PMLR, 2013. URL: `http://proceedings.mlr.press/v28/zemel13.html`.

Proof (Proof of Lemma 1). Recall that for k-center and k-supplier, speaking of the cost of an LP solution is a bit sloppy; we mean that $(\hat{x}, \hat{y})$ is a feasible solution in the LP with threshold $\hat{c}$.

The definition of $(\hat{x}, \hat{y})$ means the following. For every (fractional) assignment from a point j to a center i', we look at the cluster with center $i = \bar{\phi}(i')$ to which i' is assigned to by $\bar{\phi}$. We then 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. Notice that the (fractional) number of points assigned to i after this process may be less than one since $(\bar{x}, \bar{y})$ may include centers that are very close together.

Since that $\hat{y}$ is simply $\bar{y}$ it is integral as well and has the same number of centers, thus $\hat{y}$ also satisfies (7) if the problem uses it. Next, we observe that $(\hat{x}, \hat{y})$ satisfies fairness, i.e., respects (12) and (13). This is true because (x^{LP}, y^{LP}) satisfies them, and because we move *all* assignment from a center i' to the same center $\bar{\phi}(i')$. This shifting operation preserves the fairness. Inequality (5) is true because we only move assignments to centers that are fully open in $(\bar{x}, \bar{y})$, i.e., the inequality can not be violated as long as (4) is true (which it is for (x^{LP}, y^{LP}) since it is a feasible LP solution). Equality (4) is true for $(\hat{x}, \hat{y})$ since all assignment of j is moved to some fully open center. Thus $(\hat{x}, \hat{y})$ is a feasible solution for the LP. It remains to show that $\hat{c}$ is small enough, which depends on the objective.

k-median and k-means. We start by showing this for k-median (where the distances are a metric, i.e., $\beta = 1$ in the β-triangle inequality (1)) and k-means (where the distances are a semi-metric with $\beta = 2$). We observe that here, the cost of $(\hat{x}, \hat{y})$ is

$$\hat{c} = \sum_{j \in P} \sum_{i \in L} \hat{x}_{ij} \, \mathrm{d}(i, j) = \sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} x_{i'j}^{LP} \, \mathrm{d}(i, j).$$

Now fix $i \in L$, $i' \in \bar{C}(i)$ and $j \in P$ arbitrarily. By the β-relaxed triangle inequality, $\mathrm{d}(i, j) \leq \beta \cdot \mathrm{d}(i', j) + \beta \cdot \mathrm{d}(i', i)$. Furthermore, we know that $i' \in \bar{C}(i)$, i.e., $\bar{\phi}(i') = i$ and $\mathrm{d}(i', i) \leq \mathrm{d}(i', \bar{\phi}(j))$. We can use this to relate $\mathrm{d}(i', i)$ to the cost that j pays in $(\bar{x}, \bar{y})$:

$$\mathrm{d}(i', i) \leq \mathrm{d}(i', \bar{\phi}(j)) \leq \beta \cdot \mathrm{d}(j, i') + \beta \cdot \mathrm{d}(j, \bar{\phi}(j)).$$

Adding this up yields

$$\sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} x_{i'j}^{LP} \, \mathrm{d}(i, j)$$
$$\leq \sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} (\beta + \beta^2) x_{i'j}^{LP} \, \mathrm{d}(i', j) + \sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} \beta^2 \cdot x_{i'j}^{LP} \, \mathrm{d}(j, \bar{\phi}(j))$$
$$= (\beta + \beta^2) \cdot c^{LP} + \beta^2 \cdot \bar{c}.$$

For $\beta = 1$ (k-median), this is $2c^{LP} + \bar{c}$, for $\beta = 2$ (k-means), we get $12c^{LP} + 8\bar{c}$ due to the additional factor of 2 from Lemma 12.

Facility location. For facility location, we have to include the facility opening costs. We open the facilities that are open in $(\bar{x}, \bar{y})$, which incurs a cost of

$\sum_{i \in L} \bar{y}_i f_i$. The distance costs are the same as for k-median, so we get a total cost of

$$\sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} 2x_{i'j}^{LP} \, \mathrm{d}(i', j) + \sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} x_{i'j}^{LP} \, \mathrm{d}(j, \bar{\phi}(j)) + \sum_{i \in L} \bar{y}_i f_i \leq 2c^{LP} + \bar{c}.$$

k-center and k-supplier. For the k-center and k-supplier proof, we again fix $i \in L$, $i' \in \bar{C}(i)$ and $j \in P$ arbitrarily and use that $\mathrm{d}(i, j) \leq \mathrm{d}(i, i') + \mathrm{d}(i', j)$. Now for k-center, we know that $\mathrm{d}(i, i') \leq \bar{c}$ since $i' \in \bar{C}(i)$, and we know that $\mathrm{d}(i', j) \leq c^{LP}$ for all j where x_{ij}^{LP} is strictly positive. Thus, if $\hat{x}_{ij}$ is strictly positive, then $\mathrm{d}(i, j) \leq \bar{c} + c^{LP}$. For k-supplier, we have no guarantee that $\mathrm{d}(i, i') \leq \bar{c}$ since i' is not necessarily an input point. Instead, $i' \in \bar{C}(i)$ means that the point j' in P which is closest to i' is assigned to i by $\bar{x}$. Since j' is the closest to i' in P, we have $\mathrm{d}(i', j') \leq \mathrm{d}(i', j)$. Furthermore, since $j' \in \bar{C}(i)$, $\mathrm{d}(i, j') \leq \bar{c}$. Thus, we get for k-supplier that

$$\mathrm{d}(i, j) \leq \mathrm{d}(i, i') + \mathrm{d}(i', j) \leq \mathrm{d}(i, j') + \mathrm{d}(i', j') + \mathrm{d}(i', j) \leq \bar{c} + 2 \cdot c^{LP}.$$

Aside from our main result, we also

- propose the notion of *essentially* fair clusterings (see Definition 2) in which feasible solutions with arbitrarily many clusters exist on every feasible instance (on instances with unfavorable color distributions, only solutions with few clusters would be feasible under previous notions of fairness and this inhibits the search for hidden structures; see Section C),
- employ an extended version of our LP rounding technique to k-center/k-supplier in the *exact* fairness model where it leads to a true 5/7-approximation, i.e., one without fairness violation, in Section E,
- extend and formalize the fairlet approach from [14] for computing fair clusterings, highlighting its strengths and limitations (see Appendix G), and
- prove that the integrality gap of the natural fair ILP is unbounded for all objectives (Section C).

Algorithm 1: Black-box assignment with essentially fair solutions.

Data: An instance I of a clustering problem, and an integral solution $(\bar{x}, \bar{y})$ to I.
Result: An essentially fair solution to I with a cost as specified in Theorem 3.

---

1 Compute an optimal solution to the LP-relaxation (x^{LP}, y^{LP}) of I (see Section 1).
2 Combine $(\bar{x}, \bar{y})$ and (x^{LP}, y^{LP}) using Lemmas 1 and 2 into an essentially fair solution to I.

Algorithm 2: Black-box approximation algorithm for essentially fair clustering.

Data: An instance I of a clustering problem.
Result: An essentially fair solution to I with an approximation factor as specified in Theorem 4.

---

1 Compute an α-approximate, integral solution $(\bar{x}, \bar{y})$ to I.
2 Use Algorithm 1 with I and $(\bar{x}, \bar{y})$ as input to obtain an essentially fair solution to I with an approximation factor as specified by Theorem 4.

A   Notation

B   NP-hardness of the fair assignment problem for k-center

In this section, we reduce the Exact Cover by 3-sets to the fair assignment problem for k-center. The input to the Exact Cover by 3-sets problem is a ground set $\mathcal{U}$ of elements and a family $\mathcal{F}$ of subsets such that each set has exactly three elements from $\mathcal{U}$. The objective is to find a set cover such that each element is included in exactly one set. For example, let $\mathcal{U} = \{a, b, c, d, e, f\}, \mathcal{F} = \{A = \{a, b, c\}, B = \{b, c, d\}, C = \{d, e, f\}\}$ be an instance. The set $\{A, C\}$ is an exact cover. We call the problem of computing a cost-minimal fair assignment of points to given centers the *fair assignment problem*. It exists once for every objective listed above. Even for k-center, the fair assignment problem is NP-hard. This can be shown by a reduction from Exact Cover by 3-sets, a variant of set cover. The input is a ground set $\mathcal{U}$ of elements and a family $\mathcal{F}$ of subsets such that each set has exactly three elements from $\mathcal{U}$. The objective is to find a set cover such that each element is included in exactly one set. For example, let $\mathcal{U} = \{a, b, c, d, e, f\}, \mathcal{F} = \{A = \{a, b, c\}, B = \{b, c, d\}, C = \{d, e, f\}\}$ be an instance. The set $\{A, C\}$ is an exact cover.

For an instance $\mathcal{U}, \mathcal{F}$ of the exact cover problem, we construct an unweighted graph, which then translates to an input for the fair assignment problem for k-center by assigning distance 1 to each edge and using the resulting graph metric. The vertices consist of $\mathcal{U}, \mathcal{F}$ and two sets defined below, $\mathcal{A}$ and $\mathcal{F}$. We start by adding an edge between all $e \in \mathcal{U}$ and any $A \in \mathcal{F}$ iff $e \in A$. We assign color red to the vertices from $\mathcal{F}$ and blue to those from $\mathcal{U}$. Then we construct a set
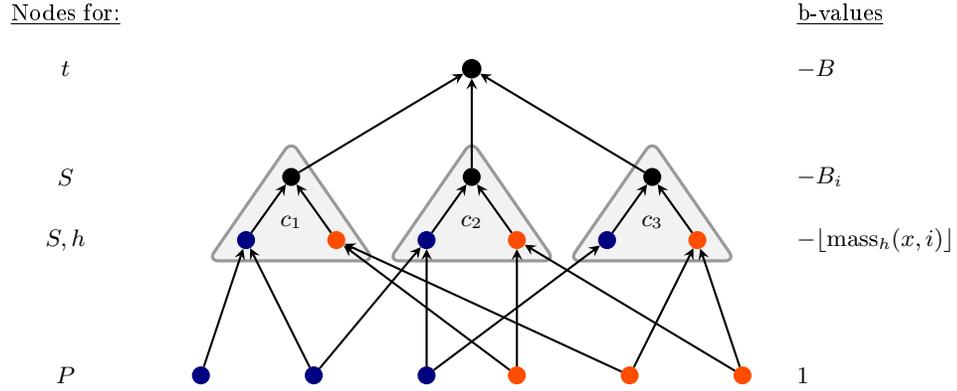
Nodes for: b-values

t $-B$

S $-B_i$

S, h $-\lfloor \mathrm{mass}_h(x, i) \rfloor$

P 1

Fig. 1. Example for the graph G used in the rounding of the x-variables.
$B_i = \lfloor \mathrm{mass}(x, i) \rfloor - \sum_{h \in Col} \lfloor \mathrm{mass}_h(x, i) \rfloor$ and $B = |P| - \sum_{i \in S} \lfloor \mathrm{mass}(x, i) \rfloor$.

$\mathcal{T}$

$\mathcal{F}$
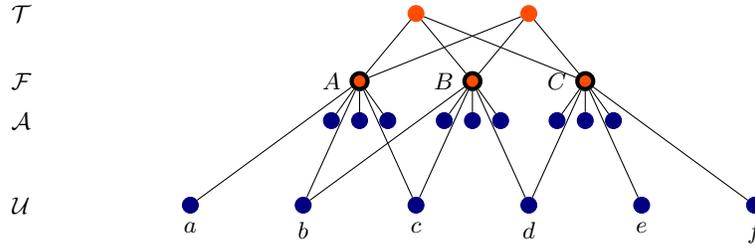
$\mathcal{A}$

$\mathcal{U}$

Fig. 2. Example for the reduction from Exact Cover with 3-sets to the fair assignment problem for k-center, with $\mathcal{U} = \{a, b, c, d, e, f\}$ and $\mathcal{F} = \{A = \{a, b, c\}, B = \{b, c, d\}, C = \{d, e, f\}\}$.

Object	Notation
the set of points	P
the set of possible locations	L
an individual point	j or p_j
an individual center	i or c_i
the number of points	n
the set of chosen locations/ the set of centers	S
an assignment of points to centers	$\phi : P \rightarrow S$
a clustering	$\mathcal{C} = (S, \phi)$
a cluster in a clustering	$P(i) = \{p \in P \mid \phi(p) = c_i\}$
the set of colors	Col
an individual color	h or col_h
color assignment	$col : P \rightarrow Col$
points with color col_h in $P' \subseteq P$	$col_h(P') = \{j \in P' \mid col(j) = col_h\}$
the amount of points with color col_h in $P' \subseteq P$	$\mathrm{mass}_h(P') = |col_h(P')|$
the amount of color col_h assigned to i in an (I)LP solution (x, y)	$\mathrm{mass}_h(x, i) = \sum_{j \in col_h(P)} x_{ij}$
the amount assigned to i in an (I)LP solution (x, y)	$\mathrm{mass}(x, i) = \sum_{j \in P} x_{ij}$
the ratio of points with color col_h in a set $P' \subseteq P$	$r_h(P') = \frac{|col_h(P')|}{|P'|}$
the number of allowed clusters	k
the distance function	$d : (P \cup L) \times (P \cup L) \rightarrow \mathbb{R}_{\geq 0}$
opening costs	f_i
relaxed boarders for the ratio	$\ell_h = \frac{p_1^h}{q_1^h}, u_h = \frac{p_2^h}{q_2^h} \in \mathbb{Q}_{\geq 0}$
a fairlet decomposition	$\mathcal{F}$
a fairlet	F_i
threshold	τ
threshold graph with threshold τ on P	$G_\tau = (P \cup L, E_\tau)$
edges in the threshold graph	$E_\tau = \{(i, j) \mid i \in L, j \in P, \mathrm{d}(i, j) \leq \tau\}$
k-center without fairness constraint (similar for other objectives)	standard k-center

Table 1. Objects and their notation

$\mathcal{A}$ which contains three auxiliary blue vertices for each vertex in $\mathcal{F}$. These are exclusively connected to their corresponding vertex in $\mathcal{F}$. Then we construct a set $\mathcal{T}$ of $|\mathcal{U}|/3$ red vertices.[8] and connect each vertex in $\mathcal{T}$ to every vertex in $\mathcal{F}$. Finally, we open a center at each vertex in $\mathcal{F}$. The construction is shown in Figure 2. Observe that the distance between an element $e \in \mathcal{U}$ and an open center at $A \in \mathcal{F}$ in this construction is 1 iff $e \in A$, and otherwise, it is 3: If $e \notin A$, then there is no edge between e and A, and since there are no direct connections between the centers, the minimum distance between e and another open center is 3.

Lemma 3. *If there exists an exact cover, there exists a fair assignment of cost 1 where the red:blue ratio is 1:3 for each cluster.*

---

[8] Note that if $|\mathcal{U}|$ is not a multiple of three, it cannot have an exact cover, so we can assume that $|\mathcal{U}|$ is a multiple of three.

Proof. Assign each red vertex $A \in \mathcal{F}$ and the three auxiliary blue vertices connected to it to the center at A. If A is in the exact cover, assign the three blue vertices representing its elements and one red vertex from $\mathcal{T}$ to the center at A. It is straightforward to verify that this assignment is fair and assigns every vertex to some center to which it is connected via a direct edge.

Lemma 4. *If there exists a fair assignment where red:blue = 1:3 for all clusters of cost less than 3, there exists an exact cover.*

Proof. For $A \in \mathcal{F}$, the red vertex at A and the three auxiliary blue vertices attached to it must be assigned to the center at A as this is the only center within distance less than 3. Also, no center can have more than two red vertices assigned to it because there are only six blue vertices in distance less than 3 of any center. Therefore, each red vertex in $\mathcal{T}$ must be assigned to a distinct center and each such center A will have exactly three blue vertices from $\mathcal{U}$ assigned to it which correspond to the elements in the set that A represents. Thus, the sets corresponding to the centers that have two red vertices assigned to them form an exact cover for $\mathcal{U}$.

C   Integrality gap of the canonical clustering LP

We show that any integral fair solution needs large clusters to implement awkward ratios of the input points. This allows us to derive a non-constant integrality gap for the canonical clustering LP.

Lemma 5. *Let P be a point set with r red and $r-1$ blue points and let $k \geq 1$. If the ratio of red points $r_{red}(C_i)$ is at most $\frac{r-k+1}{2r-2k+1}$ for each cluster C_i, then any fair solution can have at most k clusters.*

Proof. Consider a solution with $k' > k$ clusters. Since we have more red points there must be at least one cluster C_i that contains more red points than blue points. The ratio of red points $r_{red}(C_i)$ of this cluster is minimized if the solution contains $k'-1$ clusters with one blue and one red point, and one cluster with the remaining $r - k'$ blue and $r - k' + 1$ red points. However,

$$\frac{r-k'+1}{2r-2k'+1} > \frac{r-k+1}{2r-2k+1}$$

Since the highest ratio of red points in any other solution can only be higher, the claim follows.

We remark that Lemma 5 is not true for essentially fair solutions.

The canonical fair clustering ILP consists of (4)–(8) and (12)–(13). In the k-median/facility location case and in the k-means case, let write OPT_F for the optium value of its LP relaxation and and let us call the value of an optimum integral solution OPT_I. We then define the integrality gap of the ILP as $\mathrm{OPT}_I/\mathrm{OPT}_F$. In the k-center case, the ILP does not have an objective function, but we can define its integrality gap in the following sense: If τ_I, τ_F is the smallest τ such that the LP-relaxation has a feasible *integral* or *fractional* solution, respectively, then we define the integrality gap as τ_I/τ_F.

Lemma 6. *There exists a choice of non-trivial fairness intervals such that the integrality gap of the LP-relaxation of the canonical fair clustering ILP consisting of (4)–(8) and (12)–(13) is $\Omega(n)$ for the fair variants of the k-center, the k-supplier, and the facility location problem without opening costs. The integrality gap is $\Omega(n^2)$ for the fair k-means problem.*
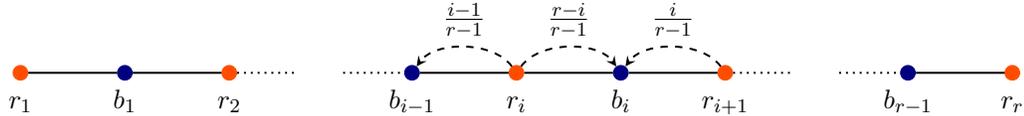

Fig. 3. Integrality gap example.

Proof. Consider the input points P lying on a line, as shown in Figure 3. Specifically, we have r red points $\{r_1, r_2, \ldots, r_r\}$ that alternate with $r-1$ blue points $\{b_1, b_2, \ldots, b_{r-1}\}$. The distance between consecutive points is 1.

We require that the ratio of the red points of each cluster is between 0 and $(r-1)/(2r-3)$ and set $k = r-1$. The input ratio $r/(2r-1)$ of the red points lies in the interior of this interval as

$$\frac{r}{2r-1} < \frac{r-1}{2r-3} \iff 2r^2 - 3r < 2r^2 - 3r + 1,$$

and thus our input is well-defined and the fairness relaxation is non-trivial. We then ask for a clustering of P with at most k centers that respects the fairness constraints.

Consider the following feasible solution for the LP-relaxation. The solution opens a center at each of the $r-1 = k$ blue points and assigns the blue point to itself and the red points on each side in the following way: for each $1 \leq i \leq r-1$, assign r_i to b_i by a fraction of $\frac{r-i}{r-1}$ and for each $2 \leq i \leq r$ assign r_i to b_{i-1} a fraction of $\frac{i-1}{r-1}$. Each red point is fully assigned in this way. We also get that in a cluster around some fixed b_i, the total assignment coming from red points is $\frac{r}{r-1}$ and the assignment coming from blue points is 1; thus, each cluster has a ratio of red points of

$$\frac{\frac{r}{r-1}}{1 + \frac{r}{r-1}} = \frac{\frac{r}{r-1}}{\frac{2r-1}{r-1}} = \frac{r}{2r-1}.$$

We therefore respect the balance requirements.

However, as $(r-1)/(2r-3) = (r-k'+1)/(2r-2k'+1)$ for $k' = 2$, by Lemma 5 any integral solution satisfying the ratio requirement can at most open two centers.

- In the k-center case, the fractional solution has a radius of 1 and the integral solution has a radius of at least $\lfloor (r-1)/2 \rfloor = \Omega(n)$. The k-center problem is a special case of the k-supplier problem; thus, the integrality gap for the k-supplier problem can only be larger.

- In the k-median case, the fractional solution has a cost of $O(n)$: The blue points incur no cost and each red point r_i contributes $(r-i)/(r-1)\cdot 1 + (i-1)/(r-1)\cdot 1 = 1$ to the objective function. Since the optimum integral solution can have at most two centers, it has to contain one cluster spanning at least $\lfloor r/2 \rfloor$ consecutive points. This incurs a cost of at least $2\cdot\sum_{j=1}^{\lfloor r/4 \rfloor -1} j = \Omega(n^2)$.
- In the facility location case, we observe that we can open at most two facilities in a fair integral solution. Hence, the analysis for the k-median case carries over (even if we set all opening costs to zero).
- In the k-means case, each red point r_i incurs a cost of $(r-i)/(r-1)\cdot 1^2 + (i-1)/(r-1)\cdot 1^2 = 1$ in the fractional solution; the blue points again incur no cost as they are chosen as centers. However, the integral solution now has a cost of at least $2\cdot\sum_{j=1}^{\lfloor r/4 \rfloor -1} j^2 = \Omega(n^3)$.

This integrality gap yields a lower bound on the quality guarantee of any LP-rounding approach for this ILP. Thus, Lemma 6 implies that no fair constant factor approximation for can be achieved by rounding the canoncial fair clustering ILP. The counterexample in 6 breaks down in the essential fairness model.

D   Further related work

Using k centers to cluster points while minimizing a certain objective function has a long history in terms of results and applications. For the k-center problem in general metric spaces, the 2-approximations developed by Gonzalez [18] and Hochbaum and Shmoys [20] were shown to be tight by Hsu and Nemhauser [21]. The k-supplier problem can be 3-approximated [20], which is also tight. Facility location can be 1.488-approximated [28], which is very close to the known APX-hardness of 1.463 for the problem [19]. For k-median, a recent breakthrough has led to a 2.765-approximation [31,11], while the best hardness result lies below two [22]. The gap between best upper and lower bound is even larger for k-means, where a 6.357-approximation is the best known [4], and the newest hardness result is marginally above 1 [8,25].

The wide applicability of k-center has given rise to numerous variants that strive to incorporate useful constraints into the solution, such as capacity constraints [10,15,23], lower bounds on the size of each cluster [3,6] or allowing for outliers [12,16]. The facility location problem has also been studied under capacity constraints [2,7,9], with uniform lower bounds [5,33], and with outliers [12]. Much less is known for k-median and k-means. True constant-factor approximations so far exist only for the outlier constraint [13,24], and are not known for the variants with capacities or lower bounds. A major problem for obtaining constant factor approximations is that the natural LP for the problem has an unbounded integrality gap, which as we will see is also true for the LP with fairness constraints. Bicriteria approximations are known that either violate the capacity constraints [27,29,30] or the cardinality constraint [1].

E True approximations for fair k-center and k-supplier

We now extend our weakly supervised rounding technique for k-center and k-supplier in the case of the exact fairness model. We replace the black-box algorithm with a specific approximation algorithm, and then achieve true approximations for the fair clustering problems by informed rounding of the LP solution.

E.1 5-Approximation Algorithm for k-center

In this section, we consider the fair k-center problem with exact preservation of ratios and without any additive fairness violation.

We give a 5-approximation for this variant. The algorithm begins by choosing a set of centers. In contrast to Section 2 we do not use an arbitrary algorithm for the standard k-center problem but specifically look for nodes in the threshold graph $G_\tau = (P, E_\tau)$ where $E_\tau = \{(i,j) \mid i \neq j \in P, d(i,j) \leq \tau\}$ that form a maximal independent set S in G_τ^2. Here G_τ^t denotes the graph on P that connects all pairs of nodes with a distance at most t in G and we denote the edge set of G_τ^t by E_τ^t. As we use the following procedure independent for each connected component of G_τ, we will in the description and the following proofs of the procedure assume that G_τ is a connected graph. The procedure uses the approach by Khuller and Sussmann [23] (procedure AssignMonarchs) to find S which ensures the following property: There exists a tree T spanning all the nodes in S and two adjacent nodes in T are exactly distance 3 apart in G_τ. The procedure begins by choosing an arbitrary vertex $r \in P$, called *root*, into S and marking every node within distance 2 of r (including itself). Until all the nodes in P are marked, it chooses an unmarked node u that is adjacent to a marked node v and marks all nodes in the distance two neighborhood of u. Observe that u is exactly at distance 3 from a node $u' \in S$ chosen earlier that caused v to get marked. Thus the run of the procedure implicitly defines the tree T over the nodes of S. In case G_τ is not a connected graph this procedure is run on each connected component and the set S has the following property: There exists a forest F such that F reduced to a connected component of G_τ is a tree T spanning all the nodes of S inside of that connected component and two adjacent nodes in T are exactly distance 3 apart in G_τ.

In the next phase, we make use of some structure that feasible solutions with exact preservation of the ratios must have.

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

We use Observation 1 and the fixed set of centers S to obtain the following adjusted LP for the fractional fair k-center problem.

$$\sum_{i \in S} x_{ij} = 1, \qquad\qquad\qquad\qquad \forall j \in P \qquad (16)$$

$$\sum_{j \in col_h(P)} x_{ij} = r_h(P) \sum_{j \in P} x_{ij} \qquad\qquad \forall i \in S \qquad (17)$$

$$\sum_{\substack{j \in col_h(P) \\ (i,j) \in E_\tau^2}} x_{ij} \geq q_h \qquad\qquad \forall i \in S, \forall h \in Col \qquad (18)$$

$$x_{ij} = 0 \qquad\qquad \forall i \in S, j \in P \text{ with } (i,j) \notin E_\tau^3 \qquad (19)$$

$$0 \leq x_{ij} \leq 1 \qquad\qquad \forall i \in S, j \in P \qquad (20)$$

Here inequality (18) ensures that each cluster contains at least q_h points of color h. Let S_{opt} be the set of centers in the optimal solution and let $\phi_{opt} : P \to S_{opt}$ be the optimal fair assignment. For the correct guess τ, every center $i \in S$ has a distinct center in S_{opt} which is at most distance one away from i in G_τ. Therefore, there exists q_h points of each color h within distance two of i. This ensures that inequality (18) is satisfiable for the right guess τ. And since, every center in S_{opt} is within distance two of some $i \in S$, there exists a fair assignment of points in P to centers in S within distance three. Thus the above LP is feasible for the right τ.

Now for the final phase, the algorithm rounds a fractional solution for the above assignment LP to an integral solution of cost at most 5τ in a procedure motivated by the LP rounding approach used by Cygan et al. in [15] for the capacitated k-center problem. Let $\beta(i)$ denote the children of node $i \in S$ in the tree T. Define quantities $\Gamma(i)$ and $\delta(i)$, $\forall i \in S$ as follows:

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

$$\delta(i) = \sum_{j \in col_1(P)} x_{ij} + \sum_{i' \in \beta(i)} \delta(i') - \Gamma(i)$$

For a leaf node i in the tree T we have $\beta(i) = \emptyset$, then $\Gamma(i)$ denotes the amount of color 1 points assigned to i rounded down to the nearest multiple of q_1, while $\delta(i)$ denotes the remaining amount. The idea is to reassign the remainder to the parent of i. Then for a non leaf i' $\Gamma(i')$ denotes the amount of color 1 points assigned to i' plus the remainder that all children of i' want to reassign to i' rounded down to the nearest multiple of q_1, while $\delta(i')$ again denotes the remainder. Since by definition of q_1 the total number of points in $col_1(P)$ must be an integer multiple of q_1, $\Gamma(r)$ also denotes the the amount of color 1 points assigned to r plus the remainder that all children of r want to reassign to r and $\delta(r) = 0$.

Also note that $\Gamma(i)$ is always a positive integer multiple of q_1 for any i, and $\delta(i)$ is always non-negative and less than q_1.

One can 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 will re-route these flows (maintaining the ratio constraints) such that $\forall i \in S, j \in col_1(P)$ x_{ij} is equal to $\Gamma(i)$ which is an integral multiple of q_1.

Lemma 7. *There exists an integral assignment of all vertices with color 1 to centers in S in G_τ^5 that assigns $\Gamma(i)$ vertices with color 1 to each center $i \in S$.*

Proof. Construct the following flow network: Take sets $col_1(P)$ and S to form a bipartite graph with an edge of capacity one between a vertex $j \in col_1(P)$ and a center $i \in S$ if and only if $(i, j) \in E_\tau^5$. Connect a source s with unit capacity edges to all vertices in $col_1(P)$ and each center $i \in S$ with capacity $\Gamma(i)$ to a sink t. We now show a feasible fractional flow of value $|col_1(P)|$ in this network. For each leaf node i in T which is not the root, assign $\Gamma(i)$ amount of color 1 flow from the total incoming color 1 flow $\sum_{j \in col_1(P)} x_{ij}$ from vertices that are at most distance three away from i in G_τ and propagate the remaining $\delta(i)$ amount of color 1 flow, coming from distance two vertices, upwards to be assigned to the parent of node i. This is always possible because by definition $\delta(i) < q_1$ and constraint (18) ensures that every center has at least q_1 amount of color 1 flow coming from distance two vertices. 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 upwards by its children) and propagate $\delta(i)$ amount of color 1 flow from distance two vertices (possible due to constraint (18)). Thus every center has $\Gamma(i)$ amount of color 1 flow passing through it and it is easy to verify that the value of the total flow in the network is $|col_1(P)|$. Since the network only has integral capacities, there exists an integral max-flow of value $|col_1(P)|$.

Lemma 8. *For any reassignment 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. Say f_1 amount of color 1 flow is reassigned from center i_1 to another center i_2. Reassign $f_h = r_h \cdot f_1/r_1$ amount of color h flow from i_1 to i_2 for each color $h \in Col \setminus \{1\}$. This is possible as constraint (17) implies that the amount of color h points assigned to i_1 must be equal to $\frac{r_h}{r_1}$ times the amount of color 1 points assigned to i_1 and f_1 must be less than the amount of color 1 points assigned to i_1. It is easy to verify that the ratios at i_1 and i_2 remain unchanged as by construction the ratio of the reassigned flows is equal to the original ratio.

From Lemmas 7 and 8 we can say that 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 must be $\Gamma(i)\frac{q_h}{q_1}$ which are integers for every center $i \in S$ and every color $h \in Col$.

Lemma 9. *There exists an integral fair assignment in G_τ^5.*

Proof. Construct a flow network for color h vertices similar to the one in lemma 7: Take sets $col_h(P)$ and S to form a bipartite graph with an edge of capacity one between a vertex $j \in col_h(P)$ and a center $i \in S$ if and only if $(i, j) \in E_\tau^5$. Connect a source s with unit capacity edges to all vertices in $col_h(P)$ and each center $i \in S$ with capacity $\Gamma(i)\frac{r_h}{r_1}$ to a sink t. The above fractional assignment in G_τ^5 gives a flow for the above network. Since the network only consists of integral demands and capacities, there is an integral max-flow which gives the assignment for the color h vertices.

Theorem 5. *There exists a 5-approximation for the fair k-center problem with exact preservation of ratios.*

Proof. Follows from Lemmas 7, 8 and 9

E.2   7-approximation for k-suppliers

We adapt the algorithm in Section E.1 to work for the k-suppliers model to give a 7-approximation for the variant with exact preservation of ratios. In the k-suppliers model, we are not allowed to open centers anywhere in P. Instead, we are provided a set L of potential locations to open centers. The procedure closely resembles the k-center algorithm: construct a bipartite threshold graph $G_\tau = (P \cup L, E_\tau)$ where $E_\tau = \{(i, j) \mid i \in L, j \in P, d(i, j) \leq \tau\}$. Choose a *root* vertex $r \in P$ into S and mark all vertices in P that are within distance two. Until all vertices in P are marked, choose an unmarked vertex $u \in P$ that is distance two away from a marked vertex and mark all vertices in the distance two neighborhood of u. Note that, since G_τ is bipartite, no two vertices in P are adjacent. The vertex u is exactly at distance four from a vertex $u' \in S$ chosen earlier. This process of selecting vertices in S defines a tree T over them with the property that adjacent vertices in T are exactly at distance four of each other in G_τ. Since we apply the procedure separately for each of the connected components of the threshold graph, we may safely assume that G_τ is connected.

Let us now temporarily open one center at each vertex in S and make the following observations for the k-suppliers case:

1. Observation 1 still holds.
2. The corresponding LP is the same as the k-center LP, except it has E_τ^4 in place of E_τ^3 in constraint (19). This ensures the feasibility of the LP since every location in L is at most distance three away from some vertex in S. (Note that in case G_τ is not connected, it can happen that some locations in L are not connected to any point and therefore more than distance three away from some vertex in S, but since they are not connected to any point we can safely ignore them, as they cannot be part of the optimal solution.)
3. Lemma 7 with G_τ^6 instead of G_τ^5 holds. The extra distance of one is introduced because the distance between a child vertex and its parent vertex in T is four instead of three.
4. Lemma 8 holds as it is and Lemma 9 holds when G_τ^5 is replaced with G_τ^6.

Thus we have a distance six fair assignment to centers in S. However, this is not a valid solution for k-suppliers as $S \subseteq P$ and we are allowed to open centers only in L. So, we move each of these temporary centers to a neighboring location in L to obtain a distance seven assignment.

F   Facts about the k-means cost function

We use some well-known facts about the k-means function when extending our results for k-median to k-means. The first one is that squared distances satisfy a relaxed triangle inequality:

Lemma 10. *It holds for all $x, y, z \in \mathbb{R}^d$ that*

$$||x - z||^2 \leq 2||x - z||^2 + 2||z - y||^2.$$

The next lemma is also a folklore statement which can be extremely useful. It implies that the best 1-means is always the centroid of a point set, and has further consequences, like Lemma 12 which we state below, a fact which is also commonly used in approximation algorithms for the k-means problem.

Lemma 11. *For any $P \subset \mathbb{R}^d$, and $z \in \mathbb{R}^d$,*

$$\sum_{x \in P} ||x - z||^2 = \sum_{x \in P} ||x - \mu(P)||^2 + |P| \cdot ||\mu(P) - z||^2,$$

where $\mu(P) = \frac{1}{|P|} \sum_{x \in P} x$ is the centroid of P.

One corollary of Lemma 11 is that the optimum cost of the best discrete solution is not much more expensive than the best choice of centers from $\mathbb{R}^d$.

Lemma 12. *Let $P \subset \mathbb{R}^d$ be a set of point in the Euclidean space, and let $C^* \subset \mathbb{R}^d$ be a set of k points that minimizes the k-means objective, i.e., it minimizes*

$$\sum_{x \in P} \min_{c \in C} ||x - c||^2$$

over all choices of $C \subset \mathbb{R}^d$ with $|C| = k$. Furthermore, let $\hat{C}$ be the set of centers that minimizes the k-means objective over all choices of $C \subset P$ with $|C| = k$, i.e., the best choice of centers from P itself. Then it holds that

$$\sum_{x \in P} \min_{c \in \hat{C}} ||x - c||^2 \leq \sum_{x \in P} \min_{c \in C^*} ||x - c||^2.$$

Thus, restricting the set of centers to the input point set increases the cost of an optimal solution by a factor of at most 2.

G   Computing fair clusterings via fairlets

G.1   Fairlets and fairlet decompositions

Chierichetti et al. [14] make an important observation: Any cluster of a fair clustering is composed of *atomic* subclusters. For instance, if we are to achieve a ratio of $2/3$ of red to blue points, then a cluster with two red and three blue cannot be divided into fair subclusters. Moreover, any fair cluster with an overall balance ratio of $2/3$ consists of an integer number of atomic subclusters – each having two red and three blue points. Figure 4 visualizes this concept. We follow [14] and call these atomic microclusters *fairlets*. The concept naturally generalizes to clusterings under relaxed fairness.
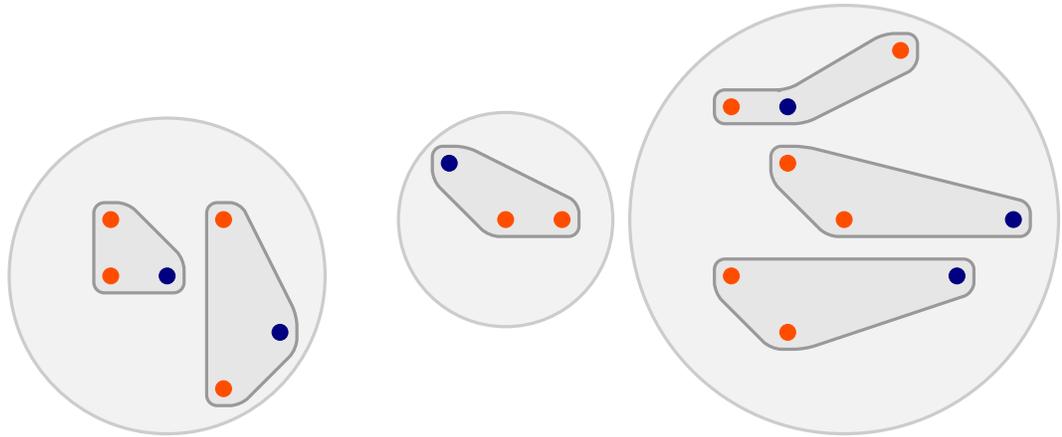

Fig. 4. A decomposition of a fair clustering into fairlets with two red points (●) and one blue point (●).

Definition 3. *A set of points $F \subseteq P$ is called a* fairlet, *if F is balanced and cannot be partitioned into multiple, non-empty balanced sets. Given $l = (l_0, \dots, l_g)$ and $u = (u_0, \dots, u_g)$ a set $F \subseteq P$ is called an l, u-fairlet, if F is l, u-balanced and can not be partitioned into multiple, non-empty l, u-balanced sets.*

When we refer to fair clusters and fairlets we also mean l, u-fair and l, u-fairlets (unless stated otherwise) as most of the proof are analogous. As observed above, every fair cluster can be partitioned into fairlets. The converse is also true: The union of disjoint fairlets is a balanced clustering.

Definition 4. *Let $P = \bigcup_{i \in I} F_i$ be a partition of P such that for each $i \in I$ the set $F_i \subseteq P$ is a fairlet (an l, u-fairlet). Then we call $\mathcal{F} = \{F_i \mid i \in I\}$ a* fairlet decomposition *(an l, u-fairlet decomposition).*

An intuitive way to look at a fairlet decomposition $\mathcal{F}$ is to view it as a fair clustering of P with $k' = |I|$ clusters (where some fairlets might share a center).[9] In the following, we denote the center of fairlet $F_i \in \mathcal{F}$ in this clustering by c_i. We then assign a cost to $\mathcal{F}$ in the following way

$$
cost(\mathcal{F}) = \begin{cases} \max_{i \in I} \max_{p \in F_i} d(p, c_i) & k-\text{center}/\text{supplier problem} \\ \sum_{i \in I} \sum_{p \in F_i} d(p, c_i) & k-\text{median problem}/\text{facility location problem} \\ \sum_{i \in I} \sum_{p \in F_i} d(p, c_i)^2 & k-\text{means problem}, \end{cases}
$$

$$
= \begin{cases} \max_{i \in I} \min_{c \in L} \max_{p \in F_i} d(p, c) & k-\text{center}/\text{supplier problem} \\ \sum_{i \in I} \min_{c \in L} \sum_{p \in F_i} d(p, c) & k-\text{median problem}/\text{facility location problem} \\ \sum_{i \in I} \min_{c \in L} \sum_{p \in F_i} d(p, c)^2 & k-\text{means problem}, \end{cases}
$$

and call the problem to find a fairlet decomposition of minimal cost the *fairlet decomposition problem*. This assignment of costs is natural: The cost of $\mathcal{F}$ is exactly the cost of the corresponding k'-clustering. It follows that the cost $\mathrm{OPT}_{\mathrm{fl}}$ of an optimum fairlet decomposition − i.e., the cost of an optimum fair clustering of P with $k' \geq k$ centers − cannot be larger than the cost OPT of an optimum fair clustering with k centers. In other words, we have $\mathrm{OPT}_{\mathrm{fl}} \leq \mathrm{OPT}$ for the fair k-center/supplier problem, for the fair k-median problem, for the fair k-means problem, and for the fair facility location problem.

G.2   The general fairlet approach for different problems

To see why fairlets help us to compute fair clusterings let us first consider a simple example. Suppose we are looking for an optimum fair k-center clustering of a point set P and denote the cost of this clustering by OPT. Also suppose that we know an optimum fairlet decomposition $\mathcal{F} := \{F_i \mid i \in I\}$ of P with cost $\mathrm{OPT}_{\mathrm{fl}}$ that assigns a center c_i to each $F_i \in \mathcal{F}$. How can we use this knowledge to find a good clustering of P? Since the union of disjoint fairlets yields a fair clustering, it seems a good idea to contract each fairlet to a single point and to then compute a colorblind clustering. This idea works indeed and was first proposed in [14]: We interpret c_i as a representative of F and compute an γ-approximate colorblind k-center clustering $\mathcal{C}'$ of the representatives $P' = \{c_i \mid i \in I\}$. Let us say that $\mathcal{C}'$ assigns c_i to a center $c(c_i) \in L$. Then, we merge the fairlets that were assigned to the same center into a fair cluster: We obtain a fair clustering $\mathcal{C}$ by assigning the points $p \in F_i$ of each fairlet $F_i \in \mathcal{F}$ to $c(c_i)$. As a union of disjoint fairlets each cluster of $\mathcal{C}$ must be fair. Also, by our previous observation that $\mathrm{OPT}_{\mathrm{fl}} \leq \mathrm{OPT}$, each fairlet must have a radius of at most OPT. Likewise, we can bound costs of the colorblind clustering $\mathcal{C}'$ by OPT: An optimum colorblind clustering cannot be more expensive than an optimum fair clustering. Thus, the radius of each cluster in $\mathcal{C}$ can at most be $(1 + \gamma) \cdot \mathrm{OPT}$. We have found a $(1 + \gamma)$-approximation!

---

[9] Obtaining a fairlet decomposition from a k'-clustering requires additional work, however: If the clustering has strictly less than k' centers, we need to subdivide some clusters into fairlets with a shared center.

In the remainder of this section, we explore what kind of approximation ratios we can expect when we cluster fairlets with a colorblind algorithm. To do so, we formalize the approach and generalize it to the k-supplier, the k-median and the k-means problem where we will need more advanced techniques for the analysis.

G.3   Using fairlets to obtain clusterings

As before, we look for clusterings of a pointset P and denote the cost of an optimum (exact or relaxed) fair k-clustering by OPT (we rely on the context to make clear whether this means k-center, k-median or k-means). We let $\mathrm{OPT}_{\mathrm{fl}}$ denote the cost of an optimum decomposition of P into k' fairlets and we denote the cost of an optimum colorblind k-clustering of P by $\mathrm{OPT}_{\mathrm{cb}}$. Since any fair k-clustering is a valid k'-clustering we have $\mathrm{OPT}_{\mathrm{fl}} \leq \mathrm{OPT}$. Likewise, any fair k-clustering is a valid colorblind k-clustering and $\mathrm{OPT}_{\mathrm{cb}} \leq \mathrm{OPT}$ follows as well; just as in the above example.

In the sequel, we suppose that we can compute an α-approximate fairlet decomposition $\mathcal{F} = \{F_i \mid i \in I\}$ of P, i.e. we suppose that $\mathrm{cost}(\mathcal{F}) \leq \alpha\mathrm{OPT}_{\mathrm{fl}} \leq \alpha\mathrm{OPT}$. We also suppose that we can compute a γ-approximate colorblind C' approximation, i.e. we ask that $\mathrm{cost}(C') \leq \gamma\mathrm{OPT}_{\mathrm{cb}} \leq \gamma\mathrm{OPT}$. The current state-of-the-art does not provide non-trivial approximations in all cases, but we postpone this issue until the next section. For now, let us assume that we have the necessary blackbox algorithms and explore the resulting guarantees of the fairlet approach.

k-center. We repeat the argument from the introductory example: Given an α-approximate fairlet decomposition of P that assigns a center c_i to each F_i in $\mathcal{F}$, we compute a γ-approximate colorblind k-center solution $\mathcal{C}'$ of $P' := \{c_i \mid i \in I\}$. If $c(c_i) \in L$ is the center assigned to c_i in $\mathcal{C}'$, we obtain a fair clustering $\mathcal{C}$ of P by assigning all $p \in F_i$ to $c(c_i)$. We can now bound the distance of any point $p \in F_i$ to its center $c(c_i)$ in $\mathcal{C}$ as $d(p, c(c_i)) \leq d(p, c_i) + d(c_i, c(c_i))$ since d is a metric. Yet, we have $d(p, c_i) \leq \alpha\mathrm{OPT}_{\mathrm{fl}} \leq \alpha\mathrm{OPT}$ and $d(c_i, c(c_i)) \leq \gamma\mathrm{OPT}_{\mathrm{cb}} \leq \gamma\mathrm{OPT}$. It follows that $\mathrm{cost}(C) \leq (\alpha+\gamma)\mathrm{OPT}$ and our algorithm is a $(\alpha+\gamma)$-approximation.

k-supplier. Given an α-approximate fairlet decomposition of P that assigns a center c_i to each F_i in $\mathcal{F}$, we cannot just compute a γ-approximate colorblind k-center solution $\mathcal{C}'$ of $P' := \{c_i \mid i \in I\}$ as we do not necessarily have $P' \subseteq P$. Instead of c_i we choose an arbitrary point $p_i \in F_i$ as a representative for F_i and compute a γ-approximate colorblind k-center solution $\mathcal{C}'$ of $P' := \{p_i \mid i \in I\}$. If $c(p_i) \in L$ is the center assigned to p_i in $\mathcal{C}'$, we obtain a fair clustering $\mathcal{C}$ of P by assigning all $p \in F_i$ to $c(p_i)$. We can now bound the distance of any point $p \in F_i$ to its center $c(p_i)$ in $\mathcal{C}$ as $d(p, c(p_i)) \leq d(p, p_i) + d(p_i, c(p_i))$ since d is a metric. Yet, we have $d(p, p_i) \leq 2\alpha\mathrm{OPT}_{\mathrm{fl}} \leq 2\alpha\mathrm{OPT}$ and $d(p_i, c(p_i)) \leq \gamma\mathrm{OPT}_{\mathrm{cb}} \leq \gamma\mathrm{OPT}$. It follows that $\mathrm{cost}(C) \leq (2\alpha + \gamma)\mathrm{OPT}$ and our algorithm is a $(2\alpha + \gamma)$-approximation. Some of the algorithms for the fairlet decomposition will already assume choose a center as one of the points. In that case it follows that $\mathrm{cost}(C) \leq (\alpha + \gamma)\mathrm{OPT}$ and our algorithm is a $(\alpha + \gamma)$-approximation.

k-median. The k-median case is more difficult: Computing representatives and repeating the above analysis yields an additive error of αOPT *for each point*, and thus at best a $2 \cdot \alpha \cdot |P|$-approximation. A different technique is needed here.

We start by computing a γ-approximate colorblind clustering $\mathcal{C}'$ on P and denote the center assigned to $p \in P$ in $\mathcal{C}'$ by $c(p)$. For all $i \in I$ and purely for the analysis we choose the center c_i of F_i in $\mathcal{F}$ as a representative of F_i. We then look for the point

$$p_i := \arg\min_{p \in F_i} \big[d(c_i, p) + d(p, c(p))\big]$$

that among all $p \in F_i$ minimizes the distance from c_i to $c(p)$ via p. This point $p_i \in P$ must have a center $c(p_i)$ in the colorblind clustering $\mathcal{C}'$ and we assign all points $p \in F_i$ to this center $c(p_i)$. In this way, we obtain a fair clustering $\mathcal{C}$. Let us first bound the cost of assigning a single point $p \in F_i$ to its center $c(p_i)$ in $\mathcal{C}$. We have

$$d(p, c(p_i)) \leq d(p, c_i) + d(c_i, p_i) + d(p_i, c(p_i))$$

By our choice of p_i, we have $d(c_i, p_i) + d(p_i, c(p_i)) \leq d(c_i, p) + d(p, c(p))$ and it follows that

$$\begin{aligned}
\text{cost}(\mathcal{C}) &= \sum_{i \in I} \sum_{p \in F_i} d(p, c(p_i)) \\
&\leq \sum_{i \in I} \sum_{p \in F_i} \big(2d(c_i, p) + d(p, c(p))\big) \\
&= 2 \cdot \sum_{i \in I} \sum_{p \in F_i} \big(d(p, c_i) + \sum_{p \in P} d(p, c(p))\big) \\
&\leq 2 \cdot \alpha\text{OPT}_{\text{fl}} + \gamma\text{OPT}_{\text{cb}}.
\end{aligned}$$

Thus, our algorithm yields a $(2\alpha + \gamma)$-approximation.

Facility location. Observe that in the above algorithm, we only open the centers from the colorblind solution; the fairlet decomposition does not incur facility opening costs. By the same reasoning as before we get

$$\begin{aligned}
\text{cost}(\mathcal{C}) &= \sum_{i \in I} \sum_{p \in F_i} d(p, c(p_i)) + \sum_{l \in L_{\text{cb}}} f_l \\
&\leq 2 \cdot \sum_{i \in I} \sum_{p \in F_i} \big(d(p, c_i) + \sum_{p \in P} d(p, c(p))\big) + \sum_{l \in L_{\text{cb}}} f_l \\
&\leq 2 \cdot \alpha\text{OPT}_{\text{fl}} + \gamma\text{OPT}_{\text{cb}}.
\end{aligned}$$

where L_{cb} denotes the set of facilities opened by the colorblind facility location solution. We obtain a $(2\alpha + \gamma)$-approximation in this case as well.

k-means. For the *k*-means problem, we assume that $P \subset \mathbb{R}^n$ and that d is the squared Euclidean distance measure. For each fairlet F_i let $\mu(F_i)$ be its centroid, i.e. $\mu(F_i) = \frac{1}{|F_i|} \sum_{p \in F_i} p$ and let us look at a colorblind γ-approximative clustering $\mathcal{C}'$ on the instance which contains $|F_i|$ copies of $\mu(F_i)$ for each $i \in I$. Without loss of generality, we assume that all of the copies of $\mu(F_i)$ will be assigned to the same center for all $i \in I$ as otherwise assigning all copies of $\mu(F_i)$ to the nearest opened center would only decrease the cost. We then obtain $\mathcal{C}$ from $\mathcal{C}'$ by assigning all points in F_i to the center c_{F_i} to which the copies of $\mu(F_i)$ were assigned in $\mathcal{C}'$. The following identity helps us to determine the cost $cost(\mathcal{C})$ of this assignment.

Proposition 1. *Given a point set $P \subseteq \mathbb{R}^d$ and a point $c \in \mathbb{R}^d$, then the 1-means cost of clustering P with c can be decomposed into*

$$\sum_{p \in P} \|p - c\|^2 = \sum_{p \in P} \|p - \mu\|^2 + |P| \cdot \|\mu - c\|^2$$

where $\mu = \frac{1}{|P|} \sum_{p \in P} p$ is known as the centroid of P.

Proposition 1 implies that the cost to assign all points $p \in F_i$ to c_{F_i} is equal to

$$\begin{aligned}
cost(\mathcal{C}) &= \sum_{i \in I} \sum_{p \in F_i} \|p - c_{F_i}\|^2 \\
&= \sum_{i \in I} \Big[\sum_{p \in F_i} \|p - \mu(F_i)]\|^2 + |F_i| \cdot \|\mu(F_i) - c_{F_i}\|^2 \Big] \\
&= \underbrace{\sum_{i \in I} \sum_{p \in F_i} \|p - \mu(F_i)\|^2}_{\leq cost(\mathcal{F})} + \underbrace{\sum_{i \in I} |F_i| \cdot \|\mu(F_i) - c_{F_i}\|^2}_{\leq cost(C')}
\end{aligned}$$

In other words, the cost of $\mathcal{C}$ is at most the joint cost of $\mathcal{F}$ and $\mathcal{C}'$. Since we know that $cost(\mathcal{F})$ is bounded by αOPT it remains to bound $cost(\mathcal{C}')$, the clustering on the centroids. To that aim, let us consider an optimum colorblind clustering of P as a reference; denote its set of centers by C. For each point $p \in P$ let $c(p)$ denote the center in C closest to p.

For any fixed $p \in F_i$ assigning all copies of $\mu(F_i)$ to $c(p)$ yields a cost of

$$|F_i| \cdot \|\mu(F_i) - c(p)\|^2 \leq |F_i| \cdot (2\|\mu(F_i) - p\|^2 + 2\|p - c(p)\|^2)$$

by the 2-relaxed triangle inequality. In each fairlet F_i, we now look at those points whose distance to $\mu(F_i)$ is at least the median of these distances accross F_i. Formally, we define for each $F_i \in \mathcal{F}$:

$$N_i := \Big\{ p \in F_i \ \Big| \ \exists \ \Big\lfloor \frac{|F_i|}{2} \Big\rfloor \text{ points } q \neq p \text{ with } \|p - \mu(F_i)\|^2 \leq \|q - \mu(F_i))\|^2 \Big\}$$

and let $\bar{p}_i := \arg\min_{p \in N_i} \|p - c(p)\|^2$ be a point with minimum distance to its center in C', among all points in N_i. With that we obtain $|F_i| \cdot \|\mu(F_i) - \bar{p}_i)\|^2 \leq$

$2 \sum_{p \in F_i} \|\mu(F_i) - p\|^2$ and $|F_i| \cdot \|\bar{p}_i - c(\bar{p}_i)\|^2 \leq 2 \sum_{p \in F_i} \|p - c(p)\|^2$ for all $i \in I$. This implies that there exists a clustering on the centroids with a cost of at most

$$\sum_{i \in I} \Big(4 \cdot \sum_{p \in F_i} \|\mu(F_i) - p\|^2 \Big) + \sum_{i \in I} \Big(4 \cdot \sum_{p \in F_i} \|p - c(q)\|^2 \Big).$$

The first sum is at most 4 times the cost of $\mathcal{F}$ and therefore bounded by $4\alpha\mathrm{OPT}$. The second sum is at most 4 times the cost of an optimal (colorblind) clustering and is therefore bounded by $4\mathrm{OPT}$. This shows that there exists a clustering on the centroids with a cost of at most $4(\alpha+1)\mathrm{OPT}$. The γ-approximation therefore has a total cost of at most $\gamma \cdot 4(\alpha + 1)\mathrm{OPT}$ and the computed clustering on P has cost of at most $(4\gamma(\alpha + 1) + \alpha)\mathrm{OPT}$. Therefore, it is a $(4\gamma(\alpha + 1) + \alpha)$-approximation.

G.4   Computing fairlet decompositions

Having reviewed fairlet based black box approximations for the different clustering objectives in the previous section, we now turn to the implementation. What concrete approximation ratios do we obtain once we fill in the black boxes with the state-of-the-art?

The happy world of k-center The colorblind variant of the k-center problem has a 2-approximation algorithm [18,20] and thus, it remains to find an α-approximative fairlet decomposition: Together with the results from the previous section, we then obtain a $(2 + \alpha)$-approximation for the fair k-center problem. All of this section is about known results, or at least algorithms that achieve the same guarantee as the known results. The prominent feature of the k-center problem is the *threshold graph*. Observe that for k-center, the optimum *cost* is a pairwise distance: It is the maximum distance between a center (which is an input point itself) and the furthest away point assigned to it. Thus, we can guess the optimum cost by iterating through all $\Theta(n^2)$ pairwise distances. 'Guessing' means that we try the $\Theta(n^2)$ values, assume in each run that our guess is the optimum and compute a solution based on this assumption. Depending on the actual algorithm, some runs may fail (indicating that we had the wrong value), and the other runs will give solutions of different quality; we can then pick the best solution, which can only be better than the solution of the run where we indeed had guessed the value correctly. This trick is used in many papers on k-center. The usual way to make use of the guessed optimum value – which we name τ – is to build a threshold graph G_τ where points are connected iff they are at distance $\leq \tau$ and then observe that two points can only be in the same optimum cluster if they are connected by a path of length two in G_τ, or, equivalently, if they are connected in G_τ^2.

Two colors. Now let's start with the case that is studied in [14], i.e., we assume that $|Col| = 2$, say $Col = \{R, B\}$. We can then easily define the *balance* of a set

Q as

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

and the balance of a clustering as the smallest balance of any cluster in it. Now the main question is what balance we want to achieve. In [14], two cases are distinguished: a) the input P has balance 1, and we want to compute a clustering which also has exactly balance 1 and b) we want a clustering with balance $1/t$, but we do not know anything about the balance of P. If $\text{balance}(P) < 1/t$, there is no feasible solution, and if $\text{balance}(P) > 1/t$, then we allow additional imbalance.

When considering the difficulty of the approximation problem, it makes sense to distinguish three different cases. For the purpose of later referencing, we state them in the following definition.

Definition 5. *For the two color case, we distinguish the following cases of a fair clustering. These are the cases where we want a clustering*

1. *with balance* $\text{balance}(P)$*, and* $\text{balance}(P) = 1$*,*
2. *with balance* $\text{balance}(P)$*, and* $\text{balance}(P) = 1/t$ *for an integer* t*,*
3. *with balance* $\text{balance}(P)$*, and* $\text{balance}(P) = s/t$ *for integers* $s \leq t$*,*
4. *with balance* $b \leq \text{balance}(P)$*, and* $b = 1/t$ *for an integer* t*, or*
5. *with balance* $b \leq \text{balance}(P)$*.*

In cases 4 and 5, the target balance b may in particular be smaller than $\text{balance}(P)$. Notice that for any finite point set, $\text{balance}(P) = s/t$ for some integers s, t, so the real restriction in case 3 is that we want to match $\text{balance}(P)$ exactly. The cases 1, 2, 3 involve *exact* fair clustering. Notice that these models force us to exactly match the proportion of red and blue points of P. Under this condition, it might not even be possible to find k clusters; thus, the optimal solution may well just consist of P. This trivial clustering, however, is always a feasible solution, so the problem is well-defined.

A first algorithm for case 1. In case 1, a good fairlet decomposition is particularly easy to obtain: It consists of a matching between the red and the blue points. Assume that $C_1, \ldots, C_k$ is the optimal fair clustering of P with radius r. Then in particular, each C_i consists of an equal number of red and blue points. Thus, there exists a bipartite matching between the red and blue points where both endpoints of a matching edge are always in the same optimum cluster.

The most intuitive idea is to recover this matching (or a better one) by considering the pairwise distances. Since we consider k-center in this paragraph, we guess the optimum cost τ as described above. Under the assumption that we guessed τ correctly, we know that any two points in the same optimum cluster are at distance at most 2τ. Thus, we just look for a perfect matching in G_τ^2. If none exists, then τ was wrong. So it suffices to find the smallest τ for which we succeed. Now define a fairlet $F = \{r_i, b_i\}$ with center r_i for each edge $e = \{r_i, b_i\}$ in the perfect matching M. Since e is part of G_τ^2, the distance between r_i and b_i – and thus the radius of F – is at most $2\tau = 2\text{OPT}$, where OPT denotes

the cost of an optimum fair clustering as before. This means that our fairlet decomposition is 2-approximate since its cost is bounded by 2OPT. In total, we obtain a 4-approximation for the fair k-center problem.

An improved algorithm for case 1. Chierichetti et al. [14] claim a 3-approximation for k-center, yet the above approach only yields a 4-approximation guarantee. We now develop an improvement. Again, partition the points into pairs of one red and one blue point from the same optimal cluster. Let r, b be such a pair. Observe that then there is a point c (a center in the optimum solution) such that $d(r, c) \leq OPT$ and $d(b, c) \leq OPT$. Thus, instead of looking at the pairwise distances between all red and blue points, we do something else. For each pair r, b, we compute the point $x = x(r, b) = \arg\min_{x \in P} \max\{d(r, x), d(b, x)\}$. Also, we set $c(r, b) = \max\{d(r, x), d(b, x)\}$. Then we know that if r and b are in the same optimum clustering, then $c(r, b) \leq OPT$. We build the threshold graph $(G'_\tau)^2$ slightly differently now by including an edge for every pair r, b with $c(r, b) \leq \tau$. A perfect matching in this graph exists iff $\tau \geq OPT$. Thus, we search for the smallest τ such that we find a perfect matching in $(G'_\tau)^2$ and then know that in this matching, every point is at distance $\leq \tau$ from his partner. We again create a fairlet $F = \{r, b\}$ for each edge $e = \{r, b\}$ in the matching; this time, however, we assign $c(r, b)$ as the center of F in the decomposition. Since now every fairlet has a radius of at most $\tau = OPT$, the decomposition is exact and we obtain a 3-approximation overall. We can apply the same approach to the fair k-supplier problem when for each pair r, b, we compute the point $x = x(r, b) = \arg\min_{x \in L} \max\{d(r, x), d(b, x)\}$ and set $c(r, b) = \max\{d(r, x), d(b, x)\}$. Since the computed decomposition is then exact and we there exist 3-approximation algorithms for the standard k-supplier problem [20], we obtain a 5-approximation.

Cases 2 and 4. In the happy k-center world, cases 2 and 4 of Definition 5 can both be 4-approximated: Chierichetti et al. [14] give a 4-approximation by using a minimum cost flow. We describe a slightly simpler algorithm which assumes that we know the majority color (or compute it as a first step). Without loss of generality, assume that blue is the majority color, i.e., at least half of the points are blue.

The difference to case 1 is that the input can no longer be broken into *pairs* of red and blue points. However, we still know that the points can be partitioned into groups of one red and $\leq t$ blue points. More precisely, we know: In any clustering satisfying the fairness constraint, a cluster C with $b(C)$ blue points must have at least $\lceil b(C)/t \rceil$ red points. Thus, we can take the optimum solution and partition each optimum cluster in subgroups of one red point together with $\leq t$ blue points. Now these subgroups are exactly the fairlets we want to find.

To obtain the partitioning, we set up a flow network built upon the threshold graph. We start by adding all red points as vertices and all blue points as vertices. We add threshold edges between red and blue points in the same manner as in the easy algorithm: For every pair r, b, we add an edge if $d(r, b) \leq 2\tau$ (i.e., if they are connected in G_τ^2). This edge gets a capacity of one. Next, we add a source and a sink. The source is connected to all red points with an edge of capacity

t: This is the maximum number of points it can facilitate. The blue points are connected to the sink by an edge of capacity 1. We then compute a maximum flow in this graph. Let n_b be the number of blue points. If the flow has a value less than n_b, we know that τ is wrong: For $\tau \geq OPT$, we know by the above argumentation that we can group the points accordingly and define a flow where every blue point is facilitated by some red point. So we compute the smallest τ for which the maximum flow has value n_b (it cannot be higher since that is the capacity of the edges going into the sink). Using the red points as centers, this step gives us fairlets with radius $\leq 2OPT$ and we obtain a 4-approximation. For k-supplier we can again use the same approach. As we already use some of the points as centers we obtain a 5-approximation.

Case 3. If balance$(P) = s/t$, then the above maximum flow idea fails since we lose the anchor node that collects the points of each fairlet. However, in the happy world of k-center, we can use capacitated clustering to fill the gap. The general idea is simple (use capacitated clustering to find fairlets), however, the execution is trickier than one might expect. The following is a summary of the approximation algorithm given in Section 4.2 in [32]. The algorithm in [32] works for multiple colors, but in this section, we only discuss two color variants.

Again, we assume that blue is the majority color, which means that $|R(P)| = n_f \cdot s$ and $|B(P)| = n_f \cdot t$ for some integers n_f, s, t with $gcd(s, t) = 1$. We consider an optimal solution. Then each cluster in this solution has $j \cdot s$ red and $j \cdot t$ blue points for $j \in \mathbb{N}$, and we can group it into fairlets with s red and t blue points. In particular, this grouping induces a uniform capacitated clustering of the red points into n_f clusters with capacity s, and a uniform capacitated clustering of the blue points into n_f clusters with capacity t, and the radius of both these clusterings is at most OPT.

There is one subtlety, though: the centers of these capacitated clusterings are not necessarily of the same color. In the optimum fair clustering, the center of each cluster may be different, sometimes red, sometimes blue. This means that the capacitated clustering problems that are induced by the optimum solution are not uniform capacitated k-center solutions, but they are uniform capacitated k-supplier solutions: For the k-supplier problem, the centers do not have to be input points, instead, the input comes with a set of points P and a set of possible center locations L as we indicated at the beginning of the preliminaries. The capacitated k-supplier problem now asks to minimize $\max_{x \in P} d(x, \phi(x))$ by choosing a set $C \subset L$, $|C| \leq k$ and an assignment of the points in P to the centers in C that respects the capacities of the centers in C.

Thus, we know that the optimum solution to the following uniform capacitated k-supplier problem costs at most OPT: Let P be the red points, let L be the union of red and blue points, and set the capacity of every center to s. Thus, we use a subroutine for this problem. In the following, we will be okay if the clustering uses soft capacities, i.e., it may open centers multiple times. A solution for this problem can only be cheaper, and we won't be using the centers as centers in the final solution anyway, so the relaxation does not hurt us.

Khuller and Sussmann [23] give a 5-approximation for the uniform soft capacitated k-center problem. Unfortunately, we need the k-supplier version. Taking a k-center solution for a k-supplier version can at most double the approximation factor; thus, we use Khuller and Sussmann's algorithm, but it gives a 10-approximation for us. The smarter way would be to *adapt* Khuller and Sussmann's algorithm to work for the k-supplier variant. We conjecture that this might give a 7-approximation instead of the 10.

We apply the algorithm to the red points and obtain clusters of size s. Now we need to assign t blue points to each red cluster. This, however, can be done by computing a matching. We construct a bipartite graph. On the red side, it contains t copies of each center, i.e., in total, it contains $t \cdot n_f = b(P)$ red points. The blue side just consists of the blue points. For the edges, we again use thresholding. For threshold τ, we add an edge between a red point and a blue point if their distance is at most τ. Now we compute the smallest τ for which the resulting graph has a perfect matching. This matching assigns one blue point to every center-copy, resulting in t blue points assigned to every red cluster.

What's the quality of the resulting solution? We know something like this: We know that in the optimum solution, there is some other clustering of the red points which is coupled with a clustering of the blue points. Now if we had the correct red clustering, the distance between a blue point and all red points in the same optimum cluster would be at most $2 \cdot \mathrm{OPT}$, leading to an overall approximation ratio of 12. We can still get a similar statement by using Hall's theorem. The details of this and a overall formally more careful variant of this proof is in Theorem 22.

This finally gives us fairlets with $s + t$ points and a corresponding center. The radius of each fairlet is at most $12 \cdot \mathrm{OPT}$ and we obtain a 14-approximation in total. For the k-supplier version we obtain a 15-approximation.

Case 5. Even in the happy k-center world no constant factor approximation algorithm is known for case 5; in particular, we do not know how to find a good fairlet decomposition. The problem here is that we do not even know that the optimal clustering can be partitioned into small balanced subsets.

Multiple colors We assume without loss of generality that $|col_0(P)| \leq |col_i(P)|$ for all $1 \leq i \leq g$. Rösner and Schmidt [32] showed that case 2 with two colors and $balance(P) = 1/t$ for an integer t can be generalized to an arbitrary number of colors with $\frac{|col_i(P)|}{|col_0(P)|} \in \mathbb{N}$ for all $1 \leq i \leq g$. We still know that the points in each optimal cluster can be partitioned into groups of one point with color col_0 and $\frac{|col_i(P)|}{|col_0(P)|}$ points with color col_i. The idea is to assign points with color $col_h \in Col \setminus \{col_0\}$, independently of points with a different color, to the points with color col_0. In the end all points connected to the same point in $col_0(P)$ build a fairlet. To do so Rösner and Schmidt [32] set up a flow network analogously to case 2: It is built upon the threshold graph for each color $col_i \in Col \setminus \{col_0\}$ which contains all points with the colors col_0 or col_i. They then chose the smallest threshold for which the corresponding networks for all colors $col_i \in Col \setminus \{col_0\}$ are successful. This again results in a fairlet decomposition with radius $\leq 2OPT$

and a 4-approximation for the fair k-center problem as well as a 5-approximation for the fair k-supplier variant.

The case 3 with $r > 1, b > 1$ can similarly be generalized to instances with arbitrary many colors [32]. Let $r = \frac{|col_0(P)|}{gcd(|col_0(P)|,...,|c_g(P)|)}$, then we know that the points in each optimal cluster can be partitioned into groups of r points with color col_0 and $\frac{|col_i(P)|}{gcd(|col_0(P)|,...,|col_g(P)|)}$ points with color col_i. The approach is then again to compute a clustering on the points with color col_0 in which every cluster contains exactly r points. Dealing with the other colors independently of each other the same approach as in case 3 can be used to match $\frac{|col_i(P)|}{gcd(|col_0(P)|,...,|col_g(P)|)}$ points with color col_i to each of these sets of r points with color col_0. Again, with the 5-approximation algorithm for the capacitated k-center problem [23] this creates a fairlet decomposition with cost of at most 12OPT and 14 and 15-approximations for fair k-center and fair k-supplier.

Theoretically, case 4 can also be generalized to a case where we require that $\frac{|col_0(P')|}{|col_i(P')|} \geq 1/t_i$ for some integer t_i for all $1 \leq i \leq g$. This way every cluster in the optimal solution can be partitioned into fair subsets out of which each contains exactly one point with color col_0 and at most t_i points of color col_i. The generalization of the approach is then identical with the generalization of case 2. However, this generalization can not be described through our ratio based problem definition with l, u-balanced clusters, as in the ratio based problem definition the points with different colors can not be treated independently of each other.

Fairlet decompositions for k-median, facility location and the k-means problem We now turn to computing fairlet decompositions for clustering objectives beside k-center and assume that we want to compute exact fairlets (i.e., not l, u-fairlets) in this section.

Instances with two colors Let us first consider the simple case where we only have two colors, i.e. $Col = \{red, blue\}$. In that case we have balance$(P) = r/b$ for two integers $r, b \in \mathbb{N}$ and $gcd(r, b) = 1$.

The case $r = b = 1$. If the balance is 1, then each fairlet consists of exactly one blue and one red point. We can therefore model the fairlet decomposition problem as a matching problem as before: We construct a complete bipartite graph in which the red and blue points make up one partition, repectively. The cost of an edge between a red point p and a blue point q is the distance between p and q. For k-median/facility location, this cost is equal to assigning both p and q to p. Assigning p and q to any other center cannot yield lower costs by the triangle inequality. In the k-means case, we compute the cost of assigning both p and q to each possible center c and then define the cost of $\{p, q\}$ as the minimum, i.e. cost$(\{p, q\}) = \min_{c \in L} \|p - c\|^2 + \|q - c\|^2$. We then compute a minimum cost perfect matching to obtain a fairlet decomposition; the cost of the matching is the cost of the decomposition. Unfortunately, this is the only case leading to a constant factor approximation algorithms so far.

The case $r = 1, b > 1$. If the balance is $1/b$ for an integer b, we can solve the fairlet decomposition problem similarly: Analogously to the k-center case, we start with the n_b many blue points and cluster them into sets with exactly b points each. This, again, could be done with an approximation algorithm for the uniform capacitated problem (i.e., k-median and k-means). Unfortunately, we are not aware of any such algorithm. Still, if we set k to n_b/b and the uniform capacity to b, then any γ-approximation to the capacitated problem has cost of at most γOPT, as observed before. We can then collect the center of each of the clusters of b blue points, and similarly to the case with $r = b = 1$ compute a perfect matching between the red points and these centers. Hall's theorem then shows that it is possible to assign one red point r to every blue cluster while guaranteeing that at least one of the blue points is in the same cluster with r in the optimal fair solution. For k-median/facility location the cost incurred by r is then at most its cost in the optimal solution, plus the cost incurred by its assigned blue point in the optimal solution, plus the cost incurred by the assigned blue point in the computed approximation. This gives us a total cost of at most $(2\alpha + 1)$OPT. For k-means the cost of the red point is then at most four times its cost in the optimal solution, plus four times the cost of its assigned blue point in the optimal solution, plus four times the cost of the assigned blue point in the computed approximation. This yields a total cost of at most $4(2\alpha + 1)$OPT.

The case $r > 1, b > 1$. We further extend the previous case. Without loss of generality we assume $b \geq r$, i.e. that blue is the majority color. Using an approximation for capacitated colorblind clustering, we again cluster the blue points into sets of size b and then create r copies of each center. We can then compute a minimum cost perfect matching between the red points and the centers.

Similar to before Hall's theorem shows that there exists an injective assignment of red points to a blue point in the same optimal fair cluster such that exactly r red points are assigned to the blue points in every blue cluster. For k-median/facility location the cost of the red point is then at most their cost in the optimal solution plus the cost of their assigned blue point in the optimal solution plus the cost of the assigned blue point in the computed approximation. This gives us a total cost of at most $(2\alpha + 1)$OPT. For k-means the cost of the red point is then again at most four times this term, which makes the total cost at most $4(2\alpha + 1)$OPT.

Instances with arbitrary many colors Let $r_i := \frac{|col_i(P)|}{gcd(|col_0(P)|,...,|col_{|Col|}(P)|)}$ be the number of points with color col_i each fairlet should have. We choose one of the colors col_i and again start by clustering the points into sets of size r_i. For each other color col_j $(i \neq j)$ we then create r_j copies of each center and compute a minimum cost perfect matching between the points with color c_j and the centers. We do this for every possible choice of color col_i and then take the best solution that we obtained. Again Hall's theorem shows that there exists an assignment of points with a color other than col_i to a point with color col_i in the same optimal fair cluster with the following property: Exactly r_j points with color col_j are assigned to the points with color col_i in every cluster, and the number of points assigned to a point with color col_i is either $\left\lfloor \frac{n - |col_i(P)|}{|col_i(P)|} \right\rfloor$

or $\left\lceil \frac{n-|col_i(P)|}{|col_i(P)|} \right\rceil$. For k-median/facility location the cost of a point with a color other than col_i is then at most its cost in the optimal solution, plus the cost of its assigned point with color col_i in the optimal solution, plus the cost of the assigned point in the computed approximation. Assuming that col_i is the color whose points have the smallest average cost in the optimal fair clustering this implies that the total cost is at most $(3\alpha + 3)\mathrm{OPT}$. For k-means the cost of a point with a color other than col_i is then at most four times its cost in the optimal solution, plus four times the cost of its assigned point with color col_i in the optimal solution, plus four times the cost of the assigned point in the computed approximation. We again assume that col_i is the color whose points have the smallest average cost in the optimal fair clustering which implies that the total cost is at most $4(3\alpha + 3)\mathrm{OPT}$.