On the Cost of Essentially Fair Clusterings

Ioana O. Bercea
School of Electrical Engineering
Tel Aviv University, Israel
ioana@cs.umd.edu

Martin Groß
School of Business and Economics
RWTH Aachen, Germany
martin.gross@mailbox.org

Samir Khuller
Department of Computer Science
Northwestern University, Evanston, USA
samir.khuller@northwestern.edu

Aounon Kumar
Department of Computer Science
University of Maryland, College Park, USA
aounon@umd.edu

Clemens Rösner
Institute of Computer Science
University of Bonn, Germany
roesner@cs.uni-bonn.de

Daniel R. Schmidt
Institute of Computer Science
University of Cologne, Germany
schmidt@informatik.uni-koeln.de

Melanie Schmidt
Institute of Computer Science
University of Bonn, Germany
melanieschmidt@uni-bonn.de

Abstract

Clustering is a fundamental tool in data mining and machine learning. It partitions points into groups (clusters) and may be used to make decisions for each point based on its group. However, this process may harm protected (minority) classes if the clustering algorithm does not adequately represent them in desirable clusters – especially if the data is already biased.

At NIPS 2017, Chierichetti et al. [18] proposed a model for *fair clustering* requiring the representation in each cluster to (approximately) preserve the global fraction of each protected class. Restricting to two protected classes, they developed both 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 is a 14-approximation for fair k-center [40].

We extend and improve the known results. Firstly, we give a 5-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. The latter approximations are achieved by a framework that takes an arbitrary existing unfair (integral) solution and a fair (fractional) LP solution and combines them into an essentially fair clustering with a weakly supervised rounding scheme. In this way, a fair clustering can be established belatedly, in a situation where the centers are already fixed.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis; Theory of computation → Facility location and clustering; Theory of computation → Rounding techniques; Theory of computation → Unsupervised learning and clustering

Keywords and phrases approximation, clustering, fairness, LP rounding

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.18

Category APPROX

Related Version `arXiv:1811.10319 [cs.DS]`

Acknowledgements The authors would like to thank Sorelle Friedler for useful discussions related to the topic of fairness. The first author's work was done while a Ph.D. student at the University of Maryland.

1 Introduction

Suppose we are to reorganize school assignments in a big city. Given a long list of children starting 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 is in fact an instance of the capacitated (metric) k-center problem. So using a k-center algorithm, you obtain a solution. However, by chance you notice an odd occurrence: One school has a huge excess of boys, while another has a surplus of girls. From previous assignment iterations, you remember that the schools prefer more balanced classes.

Thus 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.[1] This can be modeled by the following combinatorial optimization problem: Given a point set, half of the points are red, the other half is blue. 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, as proposed by Chierichetti et al. [18] in the context of maintaining fairness in *unsupervised* machine learning tasks. Their model is based on the concept of *disparate impact* [39] (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[2], 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, etc. Chierichetti et al. [18] develop a 4-approximation for a special case of fair k-center and a $(t + \sqrt{3} + \epsilon)$-approximation for one case of fair k-median.

The fair clustering model as proposed by Chierichetti et al. [18] can also be used to incorporate other aspects into our school assignment example: For example, we might want to mitigate effects of gentrification or segregation. For these use cases, we need multiple colors. Then, in each cluster, the ratio between the number of points with one specific color and the total number of points shall be in some given range. 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. One downside of this notion is that a malicious user could create an illusion of fairness by including proxy points: If we wanted to create an boy-heavy school in our above example, we could still achieve the desired parity by assigning only girls that are very unlikely to attend. Thus, instead of enforcing *equal representation* in the above sense, one could also ask for *equal opportunity* as proposed by Hardt et al. [24] for the case where we take binary decisions (i.e., $k = 2$) and have access to a labeled training set. This approach, however, raises the philosophical question if this equality of opportunity is a sufficient condition for the absence of discrimination. Rather than delving into this complex and much debated issue in this algorithmic paper, we refer to

[1] Or, incorporating the capacities, ensure that the teacher:boys:girls ratio is $1{:}\frac{s}{2}{:}\frac{s}{2}$.

[2] Imagine a point set with 49 red and 51 blue points: This cannot at all be divided into true subsets with exactly the same ratio.

the excellent surveys by Romei and Ruggieri [39] and Žliobaitė et al. [43] that systematically discuss different forms of discrimination and how they can be detected. We assume that it is the intent of the user to achieve a truly fair solution.

Finding fair clusterings turns out to be an 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. [18] 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 cannot be split further into true subsets that are also fair. 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. This is used in [40] to obtain a 14 and 15-approximation for fair k-center and k-supplier with multiple colors and exact ratio preservation.

We give an extensive overview of the existing results and further the fairlet approach in order to explore its applicability for different variants of fair clustering in the Appendix of the full version [13]. 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. Secondly, (even for k-center) it is unclear how fairlets even look like when we have multiple colors and want to allow ranges for the ratios. In this situation, subsets of very different size and composition may satisfy the desired ratio.

A different approach is to combine an LP relaxation of the constrained problem with a solution of the unconstrained problem. This approach is not specific for fair clustering; its general idea was for example used by Chakrabarty and Swamy [15] for the minimum latency facility location problem. Finding a reasonably good solution to the unconstrained problem is usually the easiest task with such an approach. Although finding a good formulation of the constrained problem as a linear program can be challenging, the main problem in such approaches is to combine the two solutions into a new solution whose cost can be bound using the quality of the two original solutions. We use such an approach. We start with a set of centers, i.e., a solution to the unconstrained problem. Then we build an LP to find a (fractional) fair solution, and use *weakly supervised LP rounding* to obtain the final integral fair solution. We use this method to prove the following statements.

▶ **Theorem 1.** *There exists a* 5 *and* 7*-approximation for the fair k-center and k-supplier problem which preserves ratios exactly.*

▶ **Theorem 2.** *Given any set of centers S, there exists an assignment ϕ' : which is essentially fair and incurs a cost that is linear in the cost S induces on the unconstrained problem and the cost of an optimal fractional fair clustering of P, for all objectives k-center, k-supplier, k-median, k-means, and facility location.*

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

Here, *essentially fair* refers to our notion of bicriteria approximation: A cluster C is *essentially fair* if there exists a fractional fair cluster C', such that for each color h the number of color h points in C differ by *at most* 1 from the mass of color h points in C'. So this is a small additive fairness violation. After the publication of our results on arXiv (Nov 2018), we have learned that in independent research, Bera et al. [12] find algorithms in a similar model as our essentially fair clustering model and achieve results similar to Corollary 3, for which they provide an almost identical analysis in their arXiv paper (Jan 2019). Theorem 1 is not affected.

We prove Theorem 2 and Corollary 3 in Section 2. Here the unconstrained starting solution can be any solution and we say our algorithm is a *black-box* approximation. We use the given integral solution to guide our rounding of a fractional solution to an LP that incorporates fairness. The proof of Theorem 1 can be found in Section 3. It is more involved as we cannot use a black-box approach, and instead need to find a suitable set of centers (a suitable integral solution) and have to adjust the weakly supervised rounding procedure.

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. All our bicriteria results work for multiple colors and approximate ratio preservation, a case for which no previous algorithm was known. As for the quality of the guarantees, compare the 4.675-approximation for essentially fair k-median clusterings with the best previously known $\Theta(t)$-approximation, which is only applicable to the case of two colors. 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.

Secondly, the black-box approach has the advantage that fairness can be established belatedly, in a situation where the centers are already given. [21, 44]. Consider our school example and notice that the location of the schools cannot be chosen. Our result says that if we are alright with essentially fair clusterings, we get a clustering which is not much more expensive than a fair clustering where the centers were chosen with the fairness constraint at hand.

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 [22] and Hochbaum and Shmoys [25] were shown to be tight by Hsu and Nemhauser [26]. The k-supplier problem can be 3-approximated [25], which is also tight. Facility location can be 1.488-approximated [35], which is very close to the known APX-hardness of 1.463 for the problem [23]. For k-median, a recent breakthrough has led to a 2.675-approximation [38, 14], while the best hardness result lies below two [27]. 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, 32].

The k-center problem allows for constant-factor approximations for many useful constraints such as capacity constraints [11, 19, 28], lower bounds on the size of each cluster [3, 6] or allowing for outliers [16, 20]. This is also true for facility location and capacities [2, 7, 10], uniform lower bounds [5, 42], and outliers [16]. Much less is known for k-median and k-means.

True constant-factor approximations so far exist only for the outlier constraint [17, 31]. A major problem for obtaining constant factor approximations is that the natural LP has an unbounded integrality gap, which is also true for the LP with fairness constraints. Bicriteria approximations are known that either violate the capacity constraints [34, 36, 37] or the cardinality constraint [1].

A clustering problem where the points have a color was considered by Li, Yi and Zhang [33]. They provided a 2-approximation for a constraint called *diversity*, which allows at most one point per color in each cluster.

The fairness constraint has been introduced by Chierichetti et al. [18]. They show a 4-approximation for the fair k-center problem with two color classes, where one color class contains t-times as many points as the other, for some integer t. Rösner and Schmidt gave a 14-approximation algorithm for k-center in the extended case with arbitrary many color classes. For the fair k-median problem with two color classes, where one color class contains t-times as many points as the other, for some integer t, Chierichetti et al. [18] also give a $\Theta(t)$-approximation. Backurs et al. [9] give an $O(d \cdot \log(n))$-approximation for a more general version of the fair k-median problem with two color classes, where a problem instance consists of n points in $\mathbb{R}^d$. For k-means the only known approximation algorithm only works for two color classes, which each contain exactly half of the points. Schmidt et al. [41] give a 32.875-approximation for this case. In parallel to our research, Bera et al. [12] have also extended the fairness model to multiple colors and approximate fairness preservation. Their model additionally allows for an overlap of the protected classes. They achieve results similar to Corollary 3.

Recent work of Kleindessner et al. [30] considers the fairness constraints in the context of spectral clustering. Fair data summarization was considered by Kleindessner et al. [29] who imposed the fairness constraint on the cluster centers alone. Specifically, they solve k-center instances with the added constraint that the chosen centers must satisfy an input distribution on the colors (i.e. out of the chosen centers, k_i must belong to color class i, where k_i is given as part of the input). While this formulation is useful for data summarization (when only the centers are reported), it is not guaranteed to lead to fair clusters overall. They propose a 5-approximation algorithm for the case of two color classes. When there are m color classes, they obtain a $(3 \cdot 2^m - 1)$-approximation.

Preliminaries

Points and locations

We are given a set of n points P and a set of potential locations L. We allow L to be infinite (when $L = \mathbb{R}^d$). The task is to open a subset $S \subseteq L$ of the locations and to assign each point in P to an open location via a mapping $\phi : P \to S$. We refer to the set of all points assigned to a location $i \in S$ by $P(i) := \phi^{-1}(i)$. The assignment incurs a cost governed by a semi-metric $d : (P \cup L) \times (P \cup L) \to \mathbb{R}_{\geq 0}$ that fulfills a *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 $\beta \geq 1$. Additionally, we may have opening costs $f_i \geq 0$ for every potential location $i \in L$ or a maximum number of centers $k \in \mathbb{N}$.

Colors and fairness

We are also given a set of *colors* $Col := \{col_1, \dots, col_g\}$, 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'. 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., $col_h(i) = col_h(P(i))$.

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 $\ell = (\ell_1, \dots, \ell_g)$ and $u = (u_1, \dots, u_g)$, if we have $r_h(P') \in [\ell_h, u_h]$ for every color $col_h \in Col$.

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 *relaxed* preservation of ratios. For the exact preservation of ratios, we ask that all clusters are exactly fair, i.e., $P(i)$ is fair for all $i \in S$.

For the relaxed preservation of ratios, we are given the lower and upper bounds $\ell = (\ell_1 = p_1^1/q_1^1, \dots, \ell_g = p_1^g/q_1^g)$ and $u = (u_1 = p_2^1/q_2^1, \dots, u_g = p_2^g/q_2^g)$ on the ratio of colors in each cluster and ask that all clusters are (ℓ, u)-*fair*. The exact case is a special case of the relaxed case where we set $\ell_h = u_h = r_h(P)$ for every color $col_h \in Col$.

Essentially fair clusterings are defined below (see Definition 6).

Objectives

We consider fair versions of several classical clustering problems. An instance is given by $I := (P, L, col, d, f, k, \ell, u)$, and our goal is to choose a solution (S, ϕ) according to one of the following objectives.

- ***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 , $L \neq P$ 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.* We show the following theorem in Section A.

▶ **Theorem 4.** *Finding an α-approximation for the fair assignment problem for k-center for $\alpha < 3$ is NP-hard.*

(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. $\phi(j) = i$. All ILP formulations have the inequalities

(2) $\sum_{i \in L} x_{ij} = 1 \; \forall j \in P$ saying that every point j is assigned to a center, the inequalities (3) $x_{ij} \leq y_i \; \forall i \in L, j \in P$ ensuring that if we assign j to i, then i must be open, and the integrality constraints (4) $y_i, x_{ij} \in \{0,1\} \; \forall i \in L, j \in P$. We may restrict the number of open centers to k with (5) $\sum_{i \in L} y_i \leq k$. 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, we ensure that points are not assigned to centers outside their range:

$$x_{ij} = 0 \qquad \text{for all } i \in L, j \in P, \{i,j\} \notin E_\tau \tag{6}$$

For the remaining clustering problems, we pick the adequate objective function from the following three (let $d_{ij} := d(i,j)$):

$$\min \sum_{i \in L, j \in P} x_{ij} d_{ij} \quad (7) \qquad \min \sum_{i \in L, j \in P} x_{ij} d_{ij}^2 \quad (8) \qquad \min \sum_{i \in L, j \in P} x_{ij} d_{ij} + \sum_{i \in L} y_i f_i \quad (9)$$

We now have all necessary constraints and objectives. For k-center and k-supplier, we use inequalities (2)-(6), 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 (2)-(5) with (7) and (8), respectively, and we get facility location by combining (2)-(4) with the objective (9). LP relaxations arise from all ILP formulations by replacing (4) 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.

$$\ell_h \sum_{j \in P} x_{ij} \leq \sum_{col(p_j) = col_h} x_{ij} \leq u_h \sum_{j \in P} x_{ij} \qquad \text{for all } i \in L, h \in Col \tag{10}$$

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 the following in Section B.

▶ **Lemma 5.** *There is a choice of non-trivial fairness intervals such that the integrality gap of the LP-relaxation of the canonical fair clustering ILP is $\Omega(n)$ for the fair k-center/k-supplier/k-median/facility location problem. The integrality gap is $\Omega(n^2)$ for the fair k-means problem.*

Essential fairness

For a point set P', $\mathrm{mass}_h(P') = |col_h(P')|$ is the *mass* of color col_h in P'. For a possibly fractional LP solution (x,y), we extend this notion to $\mathrm{mass}_h(x,i) := \sum_{j \in col_h(P)} x_{ij}$. We denote the total mass assigned to i in (x,y) by $\mathrm{mass}(x,i) = \sum_{j \in P} x_{ij}$. With this notation, we can now formalize our notion of *essential fairness*.

▶ **Definition 6** (Essential fairness). *Let I be an instance of a fair clustering problem and let (x,y) be an integral, but not necessarily fair solution to I. We say that (x,y) is* essentially fair *if there exists a fractional fair solution (x',y') for I such that $\forall i \in L$:*

$$\lfloor \mathrm{mass}_h(x',i) \rfloor \leq \mathrm{mass}_h(x,i) \leq \lceil \mathrm{mass}_h(x',i) \rceil \qquad \forall col_h \in Col \tag{11}$$

$$\textit{and} \quad \lfloor \mathrm{mass}(x',i) \rfloor \leq \mathrm{mass}(x,i) \leq \lceil \mathrm{mass}(x',i) \rceil. \tag{12}$$

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.

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. 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., inequality (10), 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., $\bar{\phi}(j) = \arg\min_{i \in S} d(j, i)$ (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.

▶ **Lemma 7.** *Let (x^{LP}, y^{LP}) and $(\bar{x}, \bar{y})$ be two solutions to the LP, where $(\bar{x}, \bar{y})$ may violate inequality* (10)*, but is integral. Then the solution defined by $\hat{y} := \bar{y}$ and*

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

satisfies inequality (10)*, $\hat{y}$ 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.*

Proof. 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 transfer 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 (5) if the problem uses it. Next, we observe that $(\hat{x}, \hat{y})$ satisfies fairness, i.e., respects (10). 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 transferring operation preserves the fairness. Inequality (3) is true because we only move assignments to centers that are fully open in $(\bar{x}, \bar{y})$, i.e., the inequality cannot be violated as long as (2) is true (which it is for (x^{LP}, y^{LP}) since it is a feasible LP solution). Equality (2) 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

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

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

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

Adding this up yields

$$\begin{aligned}
&\sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} x^{LP}_{i'j} d(i,j) \\
\leq &\sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} (\beta + \beta^2) x^{LP}_{i'j} d(i',j) + \sum_{j \in P} \sum_{i \in L} \sum_{i' \in \bar{C}(i)} \beta^2 \cdot x^{LP}_{i'j} d(j, \bar{\phi}(j)) \\
= &(\beta + \beta^2) \cdot c^{LP} + \beta^2 \cdot \bar{c}.
\end{aligned}$$

For $\beta = 1$ (k-median), this is $2c^{LP} + \bar{c}$, for $\beta = 2$ (k-means), we get $12c^{LP} + 8\bar{c}$

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

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

***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 $d(i,j) \le d(i,i') + d(i',j)$. Now for k-center, we know that $d(i,i') \le \bar{c}$ since $i' \in \bar{C}(i)$, and we know that $d(i',j) \le c^{LP}$ for all j where x^{LP}_{ij} is strictly positive. Thus, if $\hat{x}_{ij}$ is strictly positive, then $d(i,j) \le \bar{c} + c^{LP}$. For k-supplier, we have no guarantee that $d(i,i') \le \bar{c}$ 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 $d(i',j') \le d(i',j)$. Furthermore, since $j' \in \bar{C}(i)$, $d(i,j') \le \bar{c}$. Thus, we get for k-supplier that

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

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

▶ **Lemma 8.** *Let (x,y) be an α-approximate 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 $O(\mathrm{poly}(|S|+|P|))$, with $S := \{i \in L \mid y_i \ge 1\}$ 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

$$V^h := V^h_S \cup V^h_P, \quad V^h_S := \left\{v^h_i \mid i \in S\right\}, \quad V^h_P := \left\{v^h_j \mid j \in col_h(P)\right\},$$
$$A^h := \left\{(v^h_j, v^h_i) \mid i \in S, j \in col_h(P) : x_{ij} > 0\right\},$$

as well as costs c^h by $c^h_a := c_{ij}$ for $a = (v^h_j, v^h_i) \in A^h, i \in S, j \in col_h(P)$ and balances b^h by $b^h_v := 1$ if $v \in V^h_P$ and $b^h_v := -\lfloor \mathrm{mass}_h(x,i)\rfloor$ if $v = v^h_i \in V^h_S$. We use the graphs G_h to define $G = (V,A)$ by

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

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 observe that:

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 $\text{mass}_h(x,i) - \lfloor \text{mass}_h(x,i) \rfloor, \text{mass}(x,i) - \lfloor \text{mass}(x,i) \rfloor \in [0,1]$ 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 $\lfloor \text{mass}_h(x,i) \rfloor$ points of color $h \in Col$ and at least $\lfloor \text{mass}(x,i) \rfloor$ 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 $\text{mass}_h(x,i) - \lfloor \text{mass}_h(x,i) \rfloor > 0$, we have at most $\lceil \text{mass}_h(x,i) \rceil$ points of color col_h and $\lceil \text{mass}(x,i) \rceil$ 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 $\text{cost}(x) = \text{cost}(x, y)$ to the min-cost flow instances, and $\text{cost}(x) \geq \text{cost}(\hat{x})$ by construction we have $\text{cost}(\hat{x}, y) \leq \text{cost}(x, y)$.

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 7 and 8 then lead directly to Theorem 2, or, in more detail, to:

▶ **Theorem 9.** *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 given 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 known approximation factors are 2 for k-center [22, 25], 3 for k-supplier [25], 1.488 for facility location [35], 2.675 for k-median [14, 38] and 6.357 for k-means [4], which yields Corollary 3.

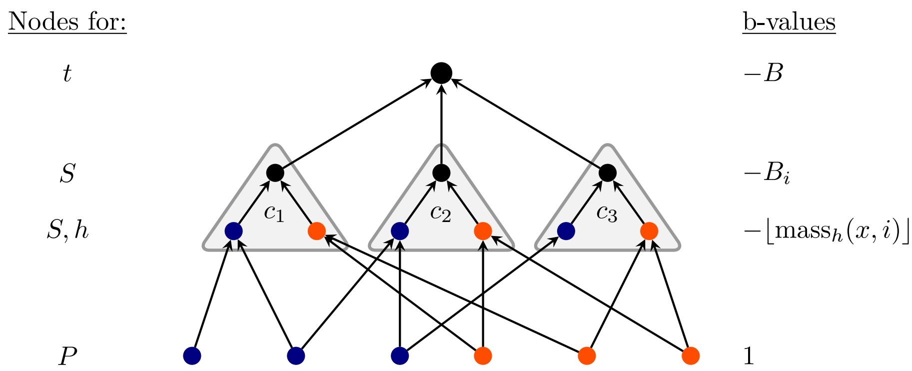

Figure 1 Example for the graph G used in the rounding of the x-variables.
$B_i = \lfloor \text{mass}(x,i) \rfloor - \sum_{h \in Col} \lfloor \text{mass}_h(x,i) \rfloor$ and $B = |P| - \sum_{i \in S} \lfloor \text{mass}(x,i) \rfloor$.

3 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.

3.1 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 which are connected by a path of length at most t in G_τ 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 [28] (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.

▶ **Observation 10.** *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 10 and the fixed set of centers S to obtain the following adjusted LP for the fractional fair k-center problem.

$$\sum_{i \in S} x_{ij} = 1, \qquad \forall j \in P \tag{13}$$

$$\sum_{j \in col_h(P)} x_{ij} = r_h(P) \sum_{j \in P} x_{ij} \qquad \forall i \in S \tag{14}$$

$$\sum_{\substack{j \in col_h(P) \\ (i,j) \in E_\tau^2}} x_{ij} \geq q_h \qquad \forall i \in S, \forall h \in Col \tag{15}$$

$$x_{ij} = 0 \qquad \forall i \in S, j \in P \text{ with } (i,j) \notin E_\tau^3 \tag{16}$$

$$0 \leq x_{ij} \leq 1 \qquad \forall i \in S, j \in P \tag{17}$$

Here inequality (15) 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 (15) 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 [19] for the capacitated k-center problem. Let $\beta(i)$ denote the children of node $i \in S$ in the tree T. Starting from the leaf nodes we recursively 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 11.** *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 (15) 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 (15)). 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 12.** *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 (14) 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 11 and 12 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 13.** *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 11: 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 14.** *There exists a 5-approximation for the fair k-center problem with exact preservation of ratios.*

Proof. Follows from Lemmas 11, 12 and 13 ◀

3.2 7-approximation for k-suppliers

We adapt the algorithm in Section 3.1 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 10 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 (16). 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 11 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 12 holds as it is and Lemma 13 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.

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 Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, and Tal Wagner. Scalable Fair Clustering. *CoRR*, abs/1902.03519, 2019. arXiv:1902.03519.

10 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.

11 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.

12 Suman K. Bera, Deeparnab Chakrabarty, and Maryam Negahbani. Fair Algorithms for Clustering. *CoRR*, abs/1901.02393, 2019. arXiv:1901.02393.

13 Ioana O. Bercea, Martin Groß, Samir Khuller, Aounon Kumar, Clemens Rösner, Daniel R. Schmidt, and Melanie Schmidt. On the cost of essentially fair clusterings. *CoRR*, abs/1811.10319, 2018. arXiv:1811.10319.

14 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.

15 Deeparnab Chakrabarty and Chaitanya Swamy. Facility location with client latencies: LP-based techniques for minimum-latency problems. *Math. Oper. Res.*, 41(3):865–883, 2016. doi:10.1287/moor.2015.0758.

16 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.

17 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.

18 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.

19 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.

20 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.

21 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.

22 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.

23 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.

24 Moritz Hardt, Eric Price, and Nati Srebro. Equality of Opportunity in Supervised Learning. In *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, pages 3315–3323, 2016.

25 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.

26 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.

27 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.

28 Samir Khuller and Yoram J. Sussmann. The Capacitated K-Center Problem. *SIAM Journal on Discrete Mathematics*, 13:403–418, 2000. doi:10.1137/S0895480197329776.

29 Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. Fair k-Center Clustering for Data Summarization. *CoRR*, abs/1901.08628, 2019. arXiv:1901.08628.

30 Matthäus Kleindessner, Samira Samadi, Pranjal Awasthi, and Jamie Morgenstern. Guarantees for Spectral Clustering with Fairness Constraints. *CoRR*, abs/1901.08668, 2019. arXiv:1901.08668.

31 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.

32 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.

33 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.

34 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`.

35 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`.

36 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`.

37 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`.

38 Shi Li and Ola Svensson. Approximating k-Median via Pseudo-Approximation. *SIAM Journal on Computing*, 45:530–547, 2016. `doi:10.1137/130938645`.

39 Andrea Romei and Salvatore Ruggieri. A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review*, 29:582–638, 2014. `doi:10.1017/S0269888913000039`.

40 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`.

41 Melanie Schmidt, Chris Schwiegelshohn, and Christian Sohler. Fair Coresets and Streaming Algorithms for Fair k-Means Clustering. *CoRR*, abs/1812.10854, 2018. `arXiv:1812.10854`.

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

43 Indrė Žliobaitė, Faisal Kamiran, and Toon Calders. Handling Conditional Discrimination. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, ICDM '11, pages 992–1001. IEEE Computer Society, 2011. `doi:10.1109/ICDM.2011.72`.

44 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`.

A 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.

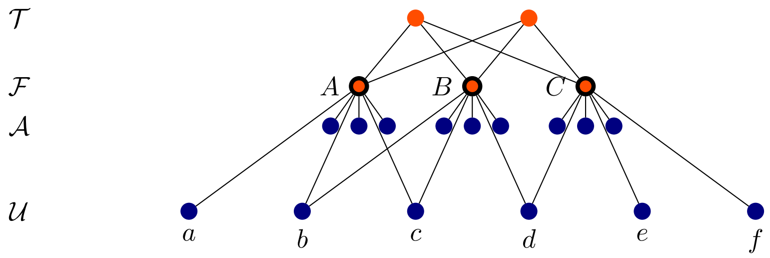

Figure 2 Example for the reduction from Exact Cover with 3-sets to the fair assignment problem for k-center, with $\mathcal{U} = \{a, b, c, d, e, f\}$ and $\mathcal{F} = \{A = \{a, b, c\}, B = \{b, c, d\}, C = \{d, e, f\}\}$.

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 $\mathcal{A}$ 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.[3] 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 15.** *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.*

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 16.** *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 $\mathcal{U}$. ◀

[3] 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.

B 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 17.** *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 17 is not true for essentially fair solutions.

The canonical fair clustering ILP consists of (2)–(6) and (10). In the k-median/facility location case and in the k-means case, let write OPT_F for the optimum 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 5.** *There is a choice of non-trivial fairness intervals such that the integrality gap of the LP-relaxation of the canonical fair clustering ILP is $\Omega(n)$ for the fair k-center/k-supplier/k-median/facility location problem. The integrality gap is $\Omega(n^2)$ for the fair k-means problem.*

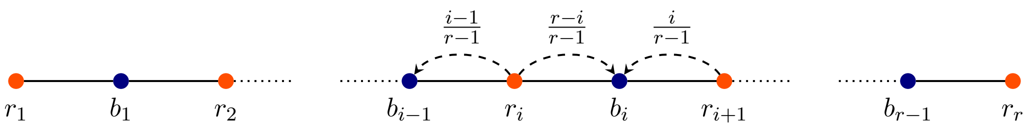

Figure 3 Integrality gap example.

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 $1 \le i \le r-1$, assign r_i to b_i by a fraction of $\frac{r-i}{r-1}$ and for each $2 \le i \le r$ 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 17 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 5 implies that no fair constant factor approximation can be achieved by rounding the canoncial fair clustering ILP. The counterexample in 5 breaks down in the essential fairness model.

C 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 18.** *It holds for all $x, y, z \in \mathbb{R}^d$ that*

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

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 20 which we state below, a fact which is also commonly used in approximation algorithms for the k-means problem.

▶ **Lemma 19.** *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 19 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 20.** *Let $P \subset \mathbb{R}^d$ be a set of point in the Euclidean space, and let $S^* \subset \mathbb{R}^d$ be a set of k points that minimizes the k-means objective, i.e., it minimizes*

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

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

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

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.