
Improved Approximation Algorithms for Individually Fair Clustering

Ali Vakilian

Toyota Technological Institute at Chicago
(TTIC)

Mustafa Yalçın

Technical University of Dortmund

Abstract

We consider the k -clustering problem with ℓ_p -norm cost, which includes k -median, k -means and k -center, under an individual notion of fairness proposed by Jung et al. [2020]: given a set of points P of size n , a set of k centers induces a *fair clustering* if every point in P has a center among its n/k closest neighbors. Mahabadi and Vakilian [2020] presented a $(p^{O(p)}, 7)$ -bicriteria approximation for fair clustering with ℓ_p -norm cost: every point finds a center within distance at most 7 times its distance to its (n/k) -th closest neighbor and the ℓ_p -norm cost of the solution is at most $p^{O(p)}$ times the cost of an optimal fair solution.

In this work, for any $\varepsilon > 0$, we present an improved $(16^p + \varepsilon, 3)$ -bicriteria for this problem. Moreover, for $p = 1$ (k -median) and $p = \infty$ (k -center), we present improved cost-approximation factors $7.081 + \varepsilon$ and $3 + \varepsilon$ respectively. To achieve our guarantees, we extend the framework of [Charikar et al., 2002, Swamy, 2016] and devise a 16^p -approximation algorithm for the facility location with ℓ_p -norm cost under matroid constraint which might be of an independent interest.

Besides, our approach suggests a reduction from our individually fair clustering to a clustering with a *group fairness requirement* proposed by Kleindessner et al. [2019], which is essentially the *median matroid* problem [Krishnaswamy et al., 2011].

1 Introduction

As automated decision-making is widely used in a diverse set of important decisions such as job hiring, loan application approval and college admission, there is a debate regarding the fairness of algorithms and machine learning methods. As there are lots of instances in which optimizing machine learning algorithms with respect to the classical measures of efficiency (e.g., accuracy, runtime and space complexity) lead to biased outputs, e.g., [Imana et al., 2021, Angwin et al., 2016], there is an extensive literature on *algorithmic fairness* which includes both *how to define the notion of fairness* and *how to design efficient algorithms with respect to fairness constraints* [Dwork et al., 2012, Chouldechova, 2017, Chouldechova and Roth, 2018, Kearns and Roth, 2019]. Clustering is one of the basic tasks in unsupervised learning and is a commonly used technique in many fields such as pattern recognition, information retrieval and data compression. Due to its wide range of applications, the clustering problem has been studied extensively under fairness consideration. Fair clustering was first introduced in a seminal work of Chierichetti et al. [2017] where they proposed the *balanced clusters* as the notion of fairness. Further, other *group fairness* notions such as *balanced centers* [Kleindessner et al., 2019] and *balanced costs* [Abbasi et al., 2021, Ghadiri et al., 2021] were also introduced as measures of fairness.

While clustering under group fairness is a well-studied domain by now, we know much less about the complexity of fairness under *individual fairness*. Motivated by the interpretation of clustering as a *facility location* problem, Jung et al. [2020] proposed an individual notion of fairness for clustering as follows: a clustering of a given pointset P is *fair* if every point in P has a center among its $(|P|/k)$ -closest neighbors. To justify, if a set of k centers are supposed to be opened, then, without any prior

knowledge, each point (or client) will expect to find a center among $1/k$ fraction of points that are closest to it. This is in particular a reasonable expectation in many scenarios. For example, people living in areas with different densities have different expectations for a “reasonably close distance”. So, while in an urban area of a major city it is reasonable for a resident to find a grocery store within a mile of her home, it is a less reasonable expectation for someone who lives in a low-density rural area. Jung et al. [2020] proves that it is NP-hard to find a fair clustering and proposed an algorithm that returns a 2-approximate fair clustering in any metric space—each point has a center at distance at most twice the distance to its $(|P|/k)$ -closest neighbor.¹

Recently, Mahabadi and Vakilian [2020] extended this notion of fairness and studied the common center-based clustering objective functions such as k -median, k -means and k -center under this individual notion of fairness. In particular, [Mahabadi and Vakilian, 2020] showed that a local search type algorithm achieves a bicriteria approximation guarantee for the aforementioned clustering problem with individual fairness. More generally, they considered the α -fair k -clustering with ℓ_p -norm cost, $\min_{C \subseteq P} \sum_{v \in P} d(v, C)^p$, and proved that a local search algorithm with swaps of size at most 4 finds a $(p^{O(p)}, 7)$ -bicriteria approximate solution: every point has a center at distance at most 7 times its “desired distance” and the ℓ_p -clustering cost of the solution is at most $p^{O(p)}$ times the optimal fair k -clustering (refer to Section 3 for more details). Given a pointset of size n and a fairness parameter $\alpha \geq 1$, for every point v , the *desired distance* of v is α times its *fair radius* where the fair radius is the distance of the (n/k) -th closest neighbor of v in the pointset.

1.1 Our Contributions

In this paper, we study the problem of α -fair k -clustering with ℓ_p -norm cost function and improve upon both fairness and cost approximation factors of the $(p^{O(p)}, 7)$ -bicriteria approximation of [Mahabadi and Vakilian, 2020] significantly. In particular, our result improves upon the $(O(\log n), 7)$ -bicriteria approximation of fair k -center clustering and achieves a $(O(1), 3)$ -bicriteria approximation.

Result 1 (restatement of Theorem 3.4) *For any $\varepsilon > 0$, $\alpha \geq 1$ and $p > 1$, there exists a $(16^p + \varepsilon, 3)$ -bicriteria approximation algorithm for α -fair k -clustering with ℓ_p -norm cost. Moreover, for*

¹Here, we assume that each point is the (first) closest neighbor to itself.

$p = 1$, which denotes the α -fair k -median problem, there exists a $(7.081 + \varepsilon, 3)$ -bicriteria approximation algorithm.

We remark that for fair k -median, our result improves upon the $(84, 7)$ -bicriteria approximation of Mahabadi and Vakilian [2020].

To achieve our approximation guarantees, we design an $e^{O(p)}$ -approximation for the problem of *facility location with ℓ_p -norm cost under matroid constraint*. This is a natural generalization of the well-known facility location problem under matroid constraint [Krishnaswamy et al., 2011, Swamy, 2016] which includes the *matroid median* problem as its special case. Our approach extends the algorithm of Swamy [2016] where we show that a careful modification of the analysis obtains the desired approximation guarantee for the more general problem of facility location with ℓ_p -norm cost. We remark that for the case of k -median ($p = 1$), we can instead employ the best-known bound for matroid-median by Krishnaswamy et al. [2018] and get $(7.081 + \varepsilon)$ -approximation.

Result 2 (restatement of Theorem A.23) *For any $p \in [1, \infty)$, there exists a 16^p -approximation algorithm for the facility location problem with ℓ_p -norm cost under matroid constraint.*

Besides our theoretical contributions, our approach essentially draws an interesting connection between the individual fairness and the group fairness notion *with balanced centers*. In particular, we show that a “density-based” decomposition of the points introduces a set of groups such that a balanced representation of them in the centers guarantees a fair solution w.r.t. the individual fairness notions considered in this paper. This observation could be of an independent interest as to the best of our knowledge is the first to connect two different notions of fairness that have been introduced for the clustering problem. Besides, the connection between our notion of individual fairness and the notion of group fairness with balanced centers has led to an improved algorithm for the fair k -center problem—we show a $(3, 3)$ -bicriteria approximation for α -fair k -center problem. Unlike our main approach, this result only holds for the k -center problem and crucially relies on properties of k -center objective function and a recent 3-approximation algorithm for k -center with balanced center [Jones et al., 2020].

Result 3 (restatement of Theorem 3.12) *For any $\varepsilon > 0$ and $\alpha \geq 1$, there exists a $(3, 3)$ -bicriteria approximation algorithm for α -fair k -center.*

1.2 Other Related Work

Clustering with group fairness constraint.

Chierichetti et al. [2017] introduced the first notion of fair clustering with balanced clusters: given a set of points coming from two distinct groups, the goal is to find a minimum *cost* clustering with proportionally balanced clusters. Their approach, which is based on a technique called *fairlet decomposition*, achieves constant factor approximations for fair k -center and k -median. Since then, several variants of clustering w.r.t. a notion of group fairness have been studied.

- **With balanced clusters.** This is the first and the most well-studied notion of fair clustering. In a series of work, this setting has been extended to address general ℓ_p -norm cost function, multiple groups, relaxed balanced requirements (with both upper and lower bound on ratio of each class in any cluster) and scalability issues [Chierichetti et al., 2017, Backurs et al., 2019, Bera et al., 2019, Bercea et al., 2019, Ahmadian et al., 2019, Schmidt et al., 2019, Huang et al., 2019].
- **With balanced centers.** Another notion of group fairness, proposed by Kleindessner et al. [2019], aims to minimize the k -center cost function and guarantee a fair representation of groups in the selected centers. Their notion is essentially k -center under partition matroid constraint. As mentioned earlier in the paper, our approach studies a generalization of this problem, facility location with ℓ_p -norm cost under matroid constraint, as a subroutine. Recently, Jones et al. [2020] designed a 3-approximation algorithm for the fair k -center with balanced centers that runs in time $O(nk)$. We remark that other clustering objective functions, in particular k -median, have been studied extensively under the partition matroid constraint, and more generally matroid constraint too [Hajiaghayi et al., 2010, Krishnaswamy et al., 2011, Charikar and Li, 2012, Chen et al., 2016, Swamy, 2016, Krishnaswamy et al., 2018]. A similar notion has been studied for the related *nearest neighbor* problem [Har-Peled and Mahabadi, 2019, Aumüller et al., 2020, 2021].
- **With balanced cost.** Recently, Abbasi et al. [2021], Ghadiri et al. [2021] independently proposed a notion of fair clustering, called *socially fair* clustering, in which the goal is to minimize the maximum cost that any group in the input pointset incurs. Makarychev and Vakilian [2021] designed an algorithm that improves upon

the $O(\ell)$ -approximation of [Abbasi et al., 2021, Ghadiri et al., 2021] for socially fair k -means and k -median and achieves $O(\log \ell / \log \log \ell)$ -approximation where ℓ denotes the number of different groups in the input. The objective of socially fair clustering was previously studied in the context of *robust* clustering [Anthony et al., 2010]. In this notion of robust algorithms, a set \mathcal{S} of possible input scenarios are provided in the input and the goal is to output a solution which is simultaneously “good” for all scenarios. Anthony et al. [2010] gave an $O(\log n + \log \ell)$ -approximation for robust k -median and a set of related problems in this model on an n -point metric space. Moreover, Bhattacharya et al. [2014] showed that it hard to approximate robust k -median by a factor better than $\Omega(\log \ell / \log \log \ell)$ unless $\text{NP} \subseteq \bigcap_{\delta > 0} \text{DTIME}(2^{n^\delta})$ which essentially shows that the approximation guarantee of Makarychev and Vakilian [2021] for socially fair k -median is tight up to a constant factor. Very recently, Chlamtáč et al. [2022] studied a more general notion of (p, q) -fair clustering which captures socially fair clustering as a special case and its objective smoothly interpolates between the objectives of k -clustering with ℓ_p -cost and socially fair clustering with ℓ_p -cost.

Inspired by the recent work on the fair allocation of public resources, Chen et al. [2019] introduced a notion of fair k -clustering as follows. Given a set of n points in a metric space, a set of k centers C is fair if no subset of n/k points $M \subset P$ have the incentive to assign themselves to a center outside C ; there is no point c' outside C such that the distance of all points in M to c' is smaller than their distance to C . Chen et al. [2019], Micha and Shah [2020] devised approximation algorithms for several variants of this notion of fair clustering.

Clustering with individual fairness constraint.

Kleindessner et al. [2020] studied a different individual notion of fairness in which a point is treated fairly, if its cluster is “stable”—the average distance of the point to its own cluster is *not larger* than the average distance of the point to the points of any other cluster. They proved that in a general metric, even deciding whether such a fair 2-clustering exists is NP-hard. Further, they showed that such fair clustering exists in one dimensional space for any values of k .

Anderson et al. [2020] proposed a distributional individual fairness for ℓ_p -norm clustering where each

point has to be mapped to a selected set of centers according to a probability distribution over the centers and then the goal is to minimize the expected ℓ_p -norm clustering cost while ensuring that “similar” points have “similar” distributional assignments to the centers.

Recently, Brubach et al. [2020, 2021] studied a *pairwise* notion of fairness for clustering which at high-level requires that the probability two points become separated be bounded by an increasing function of their pairwise distance.

Connections to priority k -center. The proposed notion of individual fairness by Jung et al. [2020] which we consider in this paper was also studied in different contexts such as *priority clustering* (or clustering with usage weights) [Plesník, 1987] and metric embedding [Chan et al., 2006, Charikar et al., 2010]. We remark that all of these results imply a 2-approximation algorithm for fair clustering. However, as in the work of Jung et al. [2020], all these result only find an approximately fair clustering and does not minimize any global clustering cost functions such as k -center, k -median and k -means.

Parallel and independent to this work, Chakrabarty and Negahbani [2021] presented a $(8, 2^{1+2/p})$ -bicriteria approximation algorithm for the individually fair k -clustering with ℓ_p -norm cost problem. In particular, their approach achieves $(8, 8)$, $(8, 4)$ and $(8, 2 + \varepsilon)$ (for arbitrarily small values of ε) for fair k -median, fair k -means and fair k -center. Table 1 provides an overview of the existing results.

Better bounds for fair k -median objective. As mentioned, we can employ the improved result of Krishnaswamy et al. [2018] and obtain $(7.081 + \varepsilon, 3)$ -approximation for fair k -median which strictly improves the recent bounds of [Chakrabarty and Negahbani, 2021]. Another improved bound related to matroid k -median is the recent result of Gupta et al. [2021]. However, we cannot apply [Gupta et al., 2021] (in a black-box manner) and get a better approximation factor. For our application (i.e., in our partition matroids, we may have $\Theta(k)$ parts) their algorithms only guarantee a *pseudo-approximation* guarantee—which assign fractional values to $O(k)$ facilities/centers—with approximation ratio $6.387 + \varepsilon$. So, while via some pre- and post-processing they can derive a true 6.387-approximation for k -median with knapsack constraint and k -median with outliers from their pseudo-approximation, their approach does not imply such approximation algorithms for the general setting with $\Theta(k)$ knapsack constraints

which we need for the case of k -median with the partition matroid constraint.

2 Preliminaries

Definition 2.1 *A distance function d satisfies the α -approximate triangle inequality over a set of points P , if $\forall u, v, w \in P, d(u, w) \leq \alpha \cdot (d(u, v) + d(v, w))$*

Observation 2.2 *Let (P, d) be a metric space. Then,*

1. *For any $\lambda > 0, p \geq 1$, The distance function $d(\cdot, \cdot)^p$ satisfies*

$$d(u, v)^p \leq (1 + \lambda)^{p-1} d(u, w)^p + \left(\frac{1 + \lambda}{\lambda} \right)^{p-1} d(w, v)^p. \quad (1)$$

In particular, for $p \geq 1$, the function $d(\cdot, \cdot)^p$ satisfies the α_p -approximate triangle inequality for $\alpha_p = 2^{p-1}$.

2. *For any $\lambda > 0, p \geq 1$, The distance function $d(\cdot, \cdot)^p$ satisfies*

$$d(u, v)^p \leq 3^{p-1} \cdot (d(u, w)^p + d(w, z)^p + d(z, v)^p). \quad (2)$$

Proof: Note that Eq. (1) holds using Lemma B.1 and the fact that $d(u, v) \leq d(u, w) + d(w, v)$. Furthermore, by setting $\lambda = 1$, $d(\cdot, \cdot)^p$ satisfies the α_p -approximate triangle inequality for $\alpha_p = 2^{p-1}$.

Next, to prove the second inequality, Eq. (2), note that $d(u, v) \leq d(u, w) + d(w, z) + d(z, v)$. Then, by an application of Lemma B.1 with $\lambda = 2$, we get Eq. (2). \square

3 A Reduction from Fair Clustering to Facility Location Under Matroid Constraint

In this section, we provide a reduction from our fair clustering problem to the problem of facility location under matroid constraint. We use P to denote the set of points in the input. We use $C \subseteq P$ to denote the subset of points that serve as centers. Throughout the paper, we consider the general ℓ_p -norm cost function which is defined as follow:

$$\text{cost}(P, C; p) := \sum_{v \in P} d(v, C)^p, \quad (3)$$

where $d(v, C)$ denotes the distance of v to its closest center in C , i.e. $d(v, C) := \min_{c \in C} d(v, c)$. This cost

| | <i>k</i> -median | | <i>k</i> -means | | <i>k</i> -center | |
|----------------------------------|-----------------------|----------|--------------------|----------|-------------------|----------|
| | Cost | Fairness | Cost | Fairness | Cost | Fairness |
| Mahabadi and Vakilian [2020] | 84 | 7 | $O(1)$ | 7 | $O(\log n)$ | 7 |
| Chakrabarty and Negahbani [2021] | 8 | 8 | 4 | 8 | $2 + \varepsilon$ | 8 |
| Ours | $7.081 + \varepsilon$ | 3 | $16 + \varepsilon$ | 3 | $3 + \varepsilon$ | 3 |

 Table 1: Comparison of the results, where $\varepsilon > 0$ is an arbitrarily small constant.

function generalizes the cost functions corresponding to *k*-median ($p = 1$), *k*-means ($p = 2$) and *k*-center ($p = \infty$).²

Next, we set up some notations to formally define the notion of fairness we consider in this paper. For every point $v \in P$, we use $B(v, r) := \{u \in P : d(v, u) \leq r\}$ to denote the subset of all points in P that are at distance at most r from v and call it the *ball* around v with radius r .

Definition 3.1 (fair radius) Let P be a set of points of size n in a metric space (X, d) and let $\ell \in [n]$ be a parameter. For every point $v \in P$, we define the fair radius $r_\ell(v)$ to be the minimum distance r such that $|B(v, r)| \geq n/\ell$. When $\ell = k$, we drop the subscript and use $r(\cdot)$ to denote $r_k(\cdot)$.

Here, we consider a more general variant of the problem studied by Mahabadi and Vakilian [2020] as follows.

Definition 3.2 (α -fair *k*-clustering) Let P be a set of points of size n in a metric space (X, d) . A set of k centers C is α -fair, if for every point $x \in P$, $d(x, C) \leq \alpha r_k(x)$.

Note that since even deciding whether a given set of points P has a fair clustering or not is NP-hard Jung et al. [2020] (i.e., $\alpha = 1$), the best we can hope for is a *bicriteria approximation* guarantee.

Definition 3.3 (bicriteria approximation)

An algorithm is a (β, γ) -bicriteria approximation for α -fair *k*-clustering w.r.t. a given ℓ_p -norm cost function if for any set of points P in the metric space (X, d) the solution SOL returned by the algorithm on P satisfies the following properties:

1. $\text{cost}(P, \text{SOL}; p) \leq \beta \cdot \text{cost}(P, \text{OPT}; p)$ where OPT denotes the optimal set of k centers for α -fair *k*-clustering of P w.r.t. the given ℓ_p -norm cost function. In particular, $\text{cost}(P, \text{OPT}; p) = \infty$ if an α -fair *k*-clustering does not exist for P .

²Note that for all $x \in \mathbb{R}^n$, $\|x\|_\infty \leq \|x\|_{\log n} \leq 2 \|x\|_\infty$. This implies that setting $p = \log n$, the objective function 2-approximates the objective of *k*-center.

2. SOL is a $(\gamma \cdot \alpha)$ -fair *k*-clustering of P .

Our main technical contribution is the following.

Theorem 3.4 For any $\varepsilon > 0$, $\alpha \geq 1$ and $p > 1$, there exists a polynomial time algorithm that computes a $(16^p + \varepsilon, 3)$ -bicriteria approximate solution for α -fair *k*-clustering with ℓ_p -norm cost. Moreover, for $p = 1$, which denotes the α -fair *k*-median problem, there exists a $(7.081 + \varepsilon, 3)$ -bicriteria approximation algorithm.

The rest of the paper is to show the above theorem. To satisfy the fairness constraint, our approach relies on the existence of a special set of regions, called *critical regions*.

Definition 3.5 (critical regions) Let P be a set of points in a metric space (X, d) and let α be the desired fairness approximation. A set of balls $\mathcal{B} = \{B(c_1, \alpha \cdot r(c_1)), \dots, B(c_m, \alpha \cdot r(c_m))\}$ where $m \leq k$ is called critical regions, if they satisfy the following properties:

1. For every $x \in P : d(x, \{c_1, \dots, c_m\}) \leq 2\alpha \cdot r(x)$
2. For any pair of centers c_i, c_j , $d(c_i, c_j) > 2\alpha \cdot \max\{r(c_i), r(c_j)\}$; in other words, critical regions are disjoint.

We now provide an algorithm that given a set of points P and a fairness parameter α , returns a set of critical regions. Our approach is similar to the methods proposed by Mahabadi and Vakilian [2020] which is a slight modification of the greedy approach of [Chan et al., 2006, Charikar et al., 2010].

Lemma 3.6 Let P be a set of points of size n in a metric space (X, d) , let k be a positive integer and let α be a parameter denoting the desired fairness guarantee. Then, Algorithm 1 returns a set of at most k critical regions.

Proof: First we show that the set of centers returned by the algorithm satisfies property (1) of the critical regions. For every points $x \in P$ let c_x denote the first center added to \mathcal{C} such that $x \in B(c_x, \alpha \cdot r(c_x))$.

Algorithm 1 outputs a set of critical regions for given parameters α, k .

-
- 1: **Input:** Fairness parameter α
 - 2: **initialize** covered points $Z \leftarrow \emptyset$, centers of the selected balls $\mathcal{C} \leftarrow \emptyset$
 - 3: **while** $Z \neq P$ **do**
 - 4: $c \leftarrow \operatorname{argmin}_{x \in P \setminus Z} r(x)$
 - 5: $\mathcal{C} \leftarrow \mathcal{C} \cup \{c\}$
 - 6: $Z \leftarrow Z \cup \{x \in P \setminus Z \mid d(x, c) \leq 2\alpha \cdot r(x)\}$
 - 7: **end while**
 - 8: **return** $\{B(c, \alpha r(c)) : c \in \mathcal{C}\}$
-

Hence, $d(x, \mathcal{C}) \leq d(x, c_x) \leq 2\alpha \cdot r(x)$ where the last inequality follows from the fact that c_x marks x as covered.

Next, consider the iteration of the algorithm in which a center c is added to \mathcal{C} . Since c is an uncovered point, its distance to any other center c' that is already in \mathcal{C} is more than $2\alpha \cdot r(c) = 2\alpha \cdot \max\{r(c), r(c')\}$ where the equality follows from the fact that centers are picked in a non-decreasing order of their fair radius. Hence, for any pair of centers in \mathcal{C} , property (2) holds.

Finally, by property (2), balls of radius $r(\cdot)$ around the centers present in \mathcal{C} are disjoint. Moreover, by the definition of fair radius, each of the balls $\{B(c, r(c))\}_{c \in \mathcal{C}}$ contains at least n/k points. Hence, the number of critical regions is at most k . \square

As shown in [Mahabadi and Vakilian, 2020], the benefit of a set of critical regions is that it reduces the problem of finding an α -fair clustering to a clustering problem with lower bound requirements, i.e., at least one center must be selected from each critical region. We say that a set of cluster centers C is *feasible* w.r.t. a set of critical regions \mathcal{B} , if for every ball $B \in \mathcal{B}$, $|C \cap B| > 0$.

Lemma 3.7 *Let $\mathcal{B} = \{B(c_1, \alpha \cdot r(c_1)), \dots, B(c_m, \alpha \cdot r(c_m))\}$ be a set of critical areas obtained from Algorithm 1 for a set of points P with parameters k and α . Then, any set of centers S that is feasible w.r.t. \mathcal{B} is (3α) -fair.*

Proof: Let S be a set of cluster centers that is feasible w.r.t. \mathcal{B} . For every point $x \in P$ let c_x denote the first center picked by Algorithm 1 such that $x \in B(c_x, \alpha \cdot r(c_x))$. Moreover, let s_x denote the center in S such that $s_x \in B(c_x, \alpha \cdot r(c_x))$. Then, for

any point $x \in P$:

$$\begin{aligned} d(x, s_x) &\leq d(x, c_x) + d(c_x, s_x) \\ &\leq 2\alpha \cdot r(x) + d(c_x, s_x) \\ &\leq 2\alpha \cdot r(x) + \alpha \cdot r(c_x) \\ &\leq 3\alpha \cdot r(x), \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from the property (1) of critical regions, the third inequality follows since $s_x \in B(c_x, \alpha \cdot r(c_x))$ and the last inequality follows since centers are added in a non-decreasing order of their fair radius in line 4 of Algorithm 1; $r(c_x) \leq r(x)$. \square

Facility location under matroid constraint.

Now we formally define the facility location problem with ℓ_p -norm cost under matroid constraint to which we reduce the problem of fair clustering with ℓ_p -norm cost. We remark that for our application, it suffices to solve the facility location under *partition matroid* constraint.

In facility location with ℓ_p -norm cost, we are given a set of facilities \mathcal{F} and a set of clients \mathcal{D} where each facility u has an opening cost of $f(u)$ and each client v is assigned with a weight (or demand) $w(v)$. The cost of assigning one unit of weight (or demand) of client v to facility u is $d(v, u)^p$. Furthermore, we are given a matroid $\mathcal{M} = (\mathcal{F}, \mathcal{I})$. Then the goal is to choose a set of facilities F that forms an independent set in \mathcal{M} and minimizes the total facility opening and client assignment cost. Formally,

$$\operatorname{argmin}_{F \in \mathcal{I}} \sum_{u \in F} f(u) + \sum_{v \in \mathcal{D}} w(v) \cdot d(v, F)^p \quad (4)$$

Next, we show a reduction from the α -fair k -clustering problem to the facility location problem under matroid constraint. Then, in Section A, we generalize the result of Swamy [2016] and devise an approximation algorithm for facility location with ℓ_p -norm cost under matroid constraint.

Reduction to facility location under matroid constraint.

Consider an instance of α -fair k -clustering on a set of points P . Let \mathcal{B} be the set of critical regions of P with parameters k and α constructed via Algorithm 1. Then, given an instance of α -fair k -clustering, Algorithm 2 constructs an instance of facility location problem under matroid constraint. Before stating the main reduction, we show that the distance function d' constructed in Algorithm 2 is a metric distance.

Lemma 3.8 *The distance function $d' : (\mathcal{F} \cup \mathcal{M}) \times (\mathcal{F} \cup \mathcal{M}) \rightarrow \mathbb{R}^+$ as constructed in Algorithm 2 constitutes a metric space.*

Algorithm 2 outputs an instance of facility location under matroid constraint corresponding to the given instance of α -fair k -clustering.

- 1: **Input:** set of points P , target number of centers k , fairness parameter α , accuracy parameter $\varepsilon < 1$, approximation guarantee of facility location under matroid constraint with ℓ_p -norm cost $\beta \geq 1$
- 2: **compute** a set of critical regions $\mathcal{B} = \{B_1, \dots, B_m\}$ via Algorithm 1 on (P, k, α)
- 3: {Construction of facilities}
- 4: **let** $P_F = \{v_f \mid v \in P\}$ be a copy of P
- 5: **let** $B_{F,i} = \{v_{f,i} \mid v \in B_i\}$ be a copy of B_i **for all** $B_i \in \mathcal{B}$
- 6: $\mathcal{F} \leftarrow P_F \cup (\bigcup_{B_i \in \mathcal{B}} B_{F,i})$ { \mathcal{F} has two distinct copies of the points that belongs to a critical ball of \mathcal{B} .}
- 7: $f(u) = 0$ **for all** $u \in \mathcal{F}$
- 8: {Construction of facilities}
- 9: **let** $\mathcal{D} = \{v_c \mid v \in P\}$ be a copy of P { \mathcal{D} is a distinct copy of P .}
- 10: $w(v) = 1$ **for all** $v \in \mathcal{D}$
- 11: {Construction of distance function $d' : (\mathcal{F} \cup \mathcal{D}) \times (\mathcal{F} \cup \mathcal{D}) \rightarrow \mathbb{R}^+$ }
- 12: **let** $\delta \leftarrow \min_{x,y \in P} d(x,y)$
- 13: **let** $d'(u,u) = 0$ **for all** $u \in \mathcal{F} \cup \mathcal{D}$
- 14: **let** $d'(v_x, u_y) = d(v,u)$ **for all** $v_x, u_y \in \mathcal{F} \cup \mathcal{D}$ where $v \neq u$
- 15: **let** $d'(v_x, v_y) = \min\{(\frac{\varepsilon(n-k)}{\beta \cdot k})^{1/p}, 1\} \cdot \delta$ **for all** $v_x, v_y \in \mathcal{F} \cup \mathcal{D}$
- 16: {Construction of matroid \mathcal{M} }
- 17: $\mathcal{M} \leftarrow$ partition matroid s.t. $|I \cap B_{F,i}| \leq 1$ **for all** $i \in [m]$ and $|I \cap P_F| \leq k - m$
- 18: **return** $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$

Theorem 3.9 *Suppose that there exists a β -approximation algorithm for the facility location with ℓ_p -norm cost under matroid constraint. Then, for any $\varepsilon > 0$, there exists a $(\beta + \varepsilon, 3)$ -bicriteria approximation for α -fair k -clustering with ℓ_p -norm cost.*

Proof: Let FACILITYMATALG be a β -approximation algorithm for facility location with ℓ_p -norm cost under matroid constraint. Consider an instance of α -fair k -clustering with ℓ_p -norm cost on pointset P and let $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$ be the instance of facility location constructed by Algorithm 2 with input parameters P, k and α . We show that the solution returned by FACILITYMATALG $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$ can be converted to a $(\beta + \varepsilon, 3)$ -bicriteria approximation for the given instance of α -fair k -clustering on P .

Let $\mathcal{B} = \{B_1, \dots, B_m\}$ be the critical regions constructed in Algorithm 2. Let SOL_F be the solution returned by FACILITYMATALG $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$ and let OPT be an optimal solution of α -fair k -clustering of P . Note that since adding centers to SOL_F only reduces the ℓ_p -cost of the solution on $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$, without loss of generality we can assume that SOL_F picks exactly one center from each of $B_{F,i}$, for $i \in [m]$, and exactly $k - m$ centers from P_F . Now we construct a solution SOL of k -clustering on P using the solution SOL_F . We start with an initially empty set of centers SOL. For each $B_i \in \mathcal{B}$, let $c_{f,i}$ denote the center in $\text{SOL}_F \cap B_{F,i}$. In the first step, we add the point $c \in P$ corresponding to $c_{f,i}$ to SOL. Next, in the second step, for each $o_f \in \text{SOL}_F \cap P_F$, we add the point $o \in P$ corresponding to o_f to SOL. Note that as some of these points may have already been added to SOL in the first step, the final solution has at most k distinct centers.

Fairness approximation. By the first step in the construction of SOL (from the given solution SOL_F), for each $i \in [m]$, $|B_i \cap \text{SOL}| \geq 1$. Hence, by Lemma 3.7, SOL is a (3α) -fair clustering of P .

Cost approximation. Note that by our reduction, all facility opening costs are set to zero. Next, we bound the assignment cost of points in P to their closest centers in SOL in terms of the the assignment cost of their corresponding client in \mathcal{D} to their closest facility in \mathcal{F} : if $v \notin \text{SOL}$, $d'(v_c, \text{SOL}_F) = d(v, \text{SOL})$; otherwise, $d'(v_c, \text{SOL}_F) = \min\{(\frac{\varepsilon(n-k)}{\beta \cdot k})^{1/p}, 1\} \cdot \delta > 0 = d(v, \text{SOL})$. Hence,

$$\text{cost}(P, \text{SOL}; p) \leq \text{cost}(\mathcal{D}, \text{SOL}_F; p). \quad (5)$$

Next, we bound the ℓ_p -cost of SOL_F on $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$ in terms of the optimal ℓ_p -cost of fair clustering P —the cost of clustering P using OPT. By the definition of α -fairness, each point $v \in P$ must have a center within distance at most $\alpha \cdot r(v)$. Hence, for each critical region $B \in \mathcal{B}$, $|\text{OPT} \cap B| \geq 1$. For each $i \in [m]$, let $c_{f,i} \in \mathcal{F}$ be the copy of an arbitrary center $c_i \in \text{OPT} \cap B_i$ in the set $B_{F,i}$. For the remaining points in OPT, we pick their corresponding copies in the set P_F ; the corresponding facilities of type c_f . Let OPT_F denote the constructed solution for the instance $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$. Since OPT_F picks exactly one point from each set $B_{F,i}$, for $i \in [m]$, and exactly $k - m$ points from P_F , OPT_F is a feasible solution of instance $(\mathcal{F}, \mathcal{D}, d', \mathcal{M})$. Moreover, since all facility opening cost are set to zero,

the ℓ_p -clustering cost of OPT on pointset P is

$$\begin{aligned}
 & \text{cost}(P, \text{OPT}; p) \\
 &= \sum_{v \in P} d(v, \text{OPT})^p \\
 &= \sum_{v \in \text{OPT}} d(v, \text{OPT})^p + \sum_{v \in P \setminus \text{OPT}} d(v, \text{OPT})^p \\
 &= \sum_{v \in \text{OPT}} (d'(v_c, \text{OPT}_F)^p - (\frac{\varepsilon(n-k)}{\beta \cdot k}) \cdot \delta^p) \\
 &\quad + \sum_{v \in P \setminus \text{OPT}} d'(v_c, \text{OPT}_F)^p \\
 &\geq \sum_{v_c \in \mathcal{D}} d'(v_c, \text{OPT}_F)^p - k \cdot (\frac{\varepsilon(n-k)}{\beta \cdot k}) \cdot \delta^p \\
 &= \text{cost}(\mathcal{D}, \text{OPT}_F; p) - \frac{\varepsilon}{\beta} (n-k) \cdot \delta^p \\
 &\geq \text{cost}(\mathcal{D}, \text{OPT}_F; p) - \frac{\varepsilon}{\beta} \cdot \text{cost}(P, \text{OPT}; p),
 \end{aligned}$$

where the last inequality holds since $\text{cost}(P, \text{OPT}; p) \geq (n-k) \cdot \delta^p$. Hence,

$$\text{cost}(\mathcal{D}, \text{OPT}_F; p) \leq (1 + \frac{\varepsilon}{\beta}) \text{cost}(P, \text{OPT}; p). \quad (6)$$

Thus,

$$\begin{aligned}
 \text{cost}(P, \text{SOL}; p) &\leq \text{cost}(\mathcal{D}, \text{SOL}_F; p) \triangleright \text{by (5)} \\
 &\leq \beta \cdot \text{cost}(\mathcal{D}, \text{OPT}_F; p) \\
 &\leq (\beta + \varepsilon) \cdot \text{cost}(P, \text{OPT}; p). \triangleright \text{by (6)}
 \end{aligned}$$

In other words, the ℓ_p -cost of clustering P using SOL is within a $\beta + \varepsilon$ factor of the optimal α -fair k -clustering of P . \square

Proof of Theorem 3.4. The proof follows from Theorem 3.9 and the 16^p -approximation algorithm of facility location with ℓ_p -norm cost under matroid constraint for $p > 1$ shown in Theorem A.23.

Next, for the case with $p = 1$, which corresponds to k -median, we can employ the approximation guarantee of [Krishnaswamy et al., 2018] and achieve a better cost approximation factor. In this setting, the proof follows from Theorem 3.9 and the 7.081-approximation algorithm of [Krishnaswamy et al., 2018] for facility location under matroid constraint. \square

3.1 A Simpler Reduction of Fair k -Center

Here, we show a reduction of the α -fair k -center problem to the k -center problem under partition matroid constraint. Then, exploiting the 3-approximation algorithm of Jones et al. [2020], we

achieve a better approximation guarantee for the α -fair k -center problem.

Consider an instance of α -fair k -center on a pointset P . Let \mathcal{B} be the set of critical regions of P with parameters k and α constructed via Algorithm 1. Then, given an instance of α -fair k -center, Algorithm 3 constructs an instance of k -center under partition matroid constraint. Similarly to Lemma 3.8,

Algorithm 3 outputs an instance of k -center under partition matroid constraint corresponding to the given instance of α -fair k -center.

- 1: **Input:** set of points P , target number of centers k , fairness parameter α , accuracy parameter $\varepsilon < 1/2$, approximation guarantee of k -center under partition matroid constraint $\beta \geq 1$
- 2: **compute** a set of critical regions $\mathcal{B} = \{B_1, \dots, B_m\}$ via Algorithm 1 on (P, k, α)
- 3: **let** $\overline{P}_0 = \{v_0 \mid v \in P\}$ be a copy of P
- 4: **let** $\overline{B}_i = \{v_i \mid v \in B_i\}$ be a copy of B_i for all $B_i \in \mathcal{B}$,
- 5: $P' \leftarrow \overline{P}_0 \cup (\bigcup_{B_i \in \mathcal{B}} \overline{B}_i)$ *{ P' has two distinct copies of the points that belongs to a critical ball of \mathcal{B} .}*
- 6: $k_i = 1$ **for all** $i \in [m]$ *{denotes that we pick at most one center from each critical ball.}*
- 7: $k_0 = k - m$
- 8: *{Construction of distance function $d' : P' \times P' \rightarrow \mathbb{R}^+$ }*
- 9: **let** $\delta \leftarrow \min_{x, y \in P} d(x, y)$
- 10: **let** $d'(u, u) = 0$ **for all** $u \in P'$
- 11: **let** $d'(v_x, u_y) = d(v, u)$ **for all** $v_x, u_y \in P'$ where $v \neq u$
- 12: **let** $d'(v_x, v_y) = \varepsilon \cdot \delta / \beta$ **for all** $v_x, v_y \in P'$
- 13: **return** $(P', \{(P_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\}, d')$

we first show that the distance function d' constructed in Algorithm 3 is a metric distance.

Lemma 3.10 *The distance function $d' : P' \times P' \rightarrow \mathbb{R}^+$ as constructed in Algorithm 3 constitutes a metric space.*

Theorem 3.11 *Suppose that there exists a β -approximation algorithm for k -center under partition matroid constraint. Then, there exists a $(\beta + \varepsilon, 3)$ -bicriteria approximation for α -fair k -center.*

Theorem 3.12 *For any $\alpha \geq 1$, there exists a polynomial time algorithm that computes a $(3 + \varepsilon, 3)$ -bicriteria approximate solution for α -fair k -center.*

Proof: The proof follows from Theorem 3.11 and the 3-approximation algorithm of [Jones et al., 2020] for k -center under partition matroid constraint. \square

Acknowledgments

AV was supported by NSF award CCF-1934843. The first author thanks Benjamin Moseley and Rudy Zhou for their helpful feedback on implications of Gupta et al. [2021] in our setting. The second author thanks Beate Bollig for her helpful comments on the exposition of the paper. Finally, the authors thank anonymous reviewers for their detailed comments which helped improve the paper.

References

- M. Abbasi, A. Bhaskara, and S. Venkatasubramanian. Fair clustering via equitable group representations. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, page 504–514, 2021.
- S. Ahmadian, A. Epasto, R. Kumar, and M. Mahdian. Clustering without over-representation. In *Proceedings of the SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 267–275, 2019.
- N. Anderson, S. K. Bera, S. Das, and Y. Liu. Distributional individual fairness in clustering. *arXiv preprint arXiv:2006.12589*, 2020.
- J. Angwin, J. Larson, S. Mattu, and L. Kirchner. Machine bias. *ProPublica*, May, 23(2016):139–159, 2016.
- B. Anthony, V. Goyal, A. Gupta, and V. Nagarajan. A plant location guide for the unsure: Approximation algorithms for min-max location problems. *Mathematics of Operations Research*, 35(1): 79–101, 2010.
- M. Aumüller, R. Pagh, and F. Silvestri. Fair near neighbor search: Independent range sampling in high dimensions. In *Proceedings of the SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 191–204, 2020.
- M. Aumüller, S. Har-Peled, S. Mahabadi, R. Pagh, and F. Silvestri. Sampling a near neighbor in high dimensions—who is the fairest of them all? *arXiv preprint arXiv:2101.10905*, 2021.
- A. Backurs, P. Indyk, K. Onak, B. Schieber, A. Vakilian, and T. Wagner. Scalable fair clustering. In *Proceedings of the International Conference on Machine Learning*, pages 405–413, 2019.
- S. Bera, D. Chakrabarty, N. Flores, and M. Negahbani. Fair algorithms for clustering. In *Advances in Neural Information Processing Systems*, pages 4955–4966, 2019.
- I. O. Bercea, M. Groß, S. Khuller, A. Kumar, C. Rösner, D. R. Schmidt, and M. Schmidt. On the cost of essentially fair clusterings. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2019.
- S. Bhattacharya, P. Chalermsook, K. Mehlhorn, and A. Neumann. New approximability results for the robust k-median problem. In *Scandinavian Workshop on Algorithm Theory*, pages 50–61, 2014.
- B. Brubach, D. Chakrabarti, J. Dickerson, S. Khuller, A. Srinivasan, and L. Tsepenekas. A pairwise fair and community-preserving approach to k-center clustering. In *International Conference on Machine Learning*, pages 1178–1189, 2020.
- B. Brubach, D. Chakrabarti, J. Dickerson, A. Srinivasan, and L. Tsepenekas. Fairness, semi-supervised learning, and more: A general framework for clustering with stochastic pairwise constraints. In *Proc. Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*, 2021.
- D. Chakrabarty and M. Negahbani. Better algorithms for individually fair k-clustering. *arXiv preprint arXiv:2106.12150*, 2021.
- T.-H. H. Chan, M. Dinitz, and A. Gupta. Spanners with slack. In *European Symposium on Algorithms*, pages 196–207, 2006.
- M. Charikar and S. Li. A dependent lp-rounding approach for the k-median problem. In *International Colloquium on Automata, Languages, and Programming*, pages 194–205, 2012.
- M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys. A constant-factor approximation algorithm for the k-median problem. *Journal of Computer and System Sciences*, 65(1):129–149, 2002.
- M. Charikar, K. Makarychev, and Y. Makarychev. Local global tradeoffs in metric embeddings. *SIAM Journal on Computing*, 39(6):2487–2512, 2010.
- D. Z. Chen, J. Li, H. Liang, and H. Wang. Matroid and knapsack center problems. *Algorithmica*, 75(1):27–52, 2016.
- X. Chen, B. Fain, L. Lyu, and K. Munagala. Proportionally fair clustering. In *International Conference on Machine Learning*, pages 1032–1041, 2019.
- F. Chierichetti, R. Kumar, S. Lattanzi, and S. Vassilvitskii. Fair clustering through fairlets. In *Advances in Neural Information Processing Systems*, pages 5036–5044, 2017.

- E. Chlamtáč, Y. Makarychev, and A. Vakilian. Approximating fair clustering with cascaded norm objectives. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2664–2683, 2022.
- A. Chouldechova. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, 5(2):153–163, 2017.
- A. Chouldechova and A. Roth. The frontiers of fairness in machine learning. *arXiv preprint arXiv:1810.08810*, 2018.
- C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. Fairness through awareness. In *Proceedings of the Innovations in Theoretical Computer Science*, pages 214–226, 2012.
- M. Ghadiri, S. Samadi, and S. Vempala. Socially fair k -means clustering. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 438–448, 2021.
- A. Gupta, B. Moseley, and R. Zhou. Structural iterative rounding for generalized k -median problems. In *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, 2021.
- M. Hajiaghayi, R. Khandekar, and G. Kortsarz. Budgeted red-blue median and its generalizations. In *Proceedings of the European Symposium on Algorithms*, pages 314–325, 2010.
- S. Har-Peled and S. Mahabadi. Near neighbor: Who is the fairest of them all? *Advances in Neural Information Processing Systems*, 32, 2019.
- L. Huang, S. Jiang, and N. Vishnoi. Coresets for clustering with fairness constraints. In *Proceedings of the Conference on Neural Information Processing Systems*, 2019.
- B. Imana, A. Korolova, and J. Heidemann. Auditing for discrimination in algorithms delivering job ads. In *Proceedings of the Web Conference 2021 (WWW '21)*, April 2021.
- M. Jones, H. Nguyen, and T. Nguyen. Fair k -centers via maximum matching. In *Proceedings of the International Conference on Machine Learning*, pages 4940–4949, 2020.
- C. Jung, S. Kannan, and N. Lutz. A center in your neighborhood: Fairness in facility location. In *Proceedings of the Symposium on Foundations of Responsible Computing*, page 5:1–5:15, 2020.
- M. Kearns and A. Roth. *The ethical algorithm: The science of socially aware algorithm design*. Oxford University Press, 2019.
- M. Kleindessner, P. Awasthi, and J. Morgenstern. Fair k -center clustering for data summarization. In *Proceedings of the International Conference on Machine Learning*, pages 3448–3457, 2019.
- M. Kleindessner, P. Awasthi, and J. Morgenstern. A notion of individual fairness for clustering. *arXiv preprint arXiv:2006.04960*, 2020.
- R. Krishnaswamy, A. Kumar, V. Nagarajan, Y. Sabharwal, and B. Saha. The matroid median problem. In *Proceedings of the Symposium on Discrete Algorithms*, pages 1117–1130, 2011.
- R. Krishnaswamy, S. Li, and S. Sandeep. Constant approximation for k -median and k -means with outliers via iterative rounding. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 646–659, 2018.
- S. Mahabadi and A. Vakilian. Individual fairness for k -clustering. In *Proceedings of the International Conference on Machine Learning*, pages 6586–6596, 2020.
- K. Makarychev, Y. Makarychev, and I. Razenshteyn. Performance of Johnson-Lindenstrauss transform for k -means and k -medians clustering. In *Proceedings of the Symposium on Theory of Computing*, pages 1027–1038, 2019.
- Y. Makarychev and A. Vakilian. Approximation algorithms for socially fair clustering. *arXiv preprint arXiv:2103.02512*, 2021.
- E. Micha and N. Shah. Proportionally fair clustering revisited. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, 2020.
- J. Plesník. A heuristic for the p -center problems in graphs. *Discrete Applied Mathematics*, 17(3): 263–268, 1987.
- M. Schmidt, C. Schwiegelshohn, and C. Sohler. Fair coresets and streaming algorithms for fair k -means. In *Proceedings of the International Workshop on Approximation and Online Algorithms*, pages 232–251, 2019.
- C. Swamy. Improved approximation algorithms for matroid and knapsack median problems and applications. *ACM Transactions on Algorithms (TALG)*, 12(4):1–22, 2016.

A $e^{O(p)}$ -Approximation for Matroid Facility Location with ℓ_p -Cost

We consider a natural LP-relaxation for the facility location problem under matroid constraint considered by Swamy [2016]. This relaxation is a generalization of the standard LP relaxation of k -median and k -means clustering (e.g. [Charikar et al., 2002]). For every facility $u \in \mathcal{F}$, the variable y_u denotes whether the facility u is open or not; $y_u = 1$ if u is open and $y_u = 0$ otherwise. For every client $v \in P$ and facility $u \in \mathcal{F}$, we have a variable x_{vu} that denotes whether u is the closest facility among the opened facilities to v . We also assume that we are given a function $w : P \rightarrow \mathbb{R}$ which denotes the demand of clients. Finally, we use r to denote the rank function of the matroid $\mathcal{M} = (\mathcal{F}, \mathcal{I})$.

LP Relaxation: FACILITYMATLP(w, \mathcal{M})

$$\begin{aligned} \text{minimize} \quad & \sum_{u \in \mathcal{F}} f(u) \cdot y_u + \sum_{v \in P, u \in \mathcal{F}} w(v) \cdot d(v, u)^p \cdot x_{vu} \\ \text{s.t.} \quad & \sum_{u \in \mathcal{F}} x_{vu} \geq 1 \quad \forall v \in P \end{aligned} \tag{7}$$

$$\sum_{u \in S} y_u \leq r(S) \quad \forall S \subseteq \mathcal{F} \tag{8}$$

$$0 \leq x_{uv} \leq y_u \quad \forall u \in \mathcal{F}, v \in P \tag{9}$$

Note that for an optimal solution (x, y) of FACILITYMATLP, may assume that for every client v , $\sum_{u \in \mathcal{F}} x_{vu} = 1$.

We follow the framework of [Charikar et al., 2002, Swamy, 2016]. First, we reduce the instance into a *well-separated* instance and then we find a half-integral solution of the well-separated instance. We start with an optimal *fractional* solution (x^*, y^*) to FACILITYMATLP(w, \mathcal{M}). The cost of this optimal fractional solution is denoted by z^* . The ultimate goal is to obtain a good integral solution whose cost is comparable to z^* . First we construct a modified instance, called *well-separated* instance. A key property of the well-separated instance is that the distance of any pair of clients v and u in this instance is “large” compared to the contributions of each of u and v w.r.t. (x^*, y^*) . Then, we prove two important statements: (1) there exists a fractional solution of the well-separated instance whose cost is not more than z^* and (2) an integral solution of the well-separated instance with cost z' can be transformed into a solution of cost $e^{O(p)}(z' + z^*)$ on the original instance.

Hence, it suffices to find an *integral* solution F' of the well-separated instance whose cost is not “much larger” than z^* . If we achieve that, then we get a “good” approximate integral solution of the original instance.

In the first step, we construct a half-integral solution \hat{y} to the well-separated instance whose cost is not more than $3^p \cdot z^*$. First, we show that there exists a solution y' with “certain structure” whose cost is at most $3^p \cdot z^*$. Then, we consider a modified cost function that plays a role as a proxy for the actual cost and show that under the new cost function we can always find a feasible half-integral solution with minimum cost. Lastly, we show that the actual cost of the constructed half-integral solution is also at most $3^p \cdot z^*$.

In the second step, we construct an integral solution \tilde{y} to the well-separated instance from the half-integral solution \hat{y} and show that the cost of the integral solution is at most $e^{O(p)}$ times the cost of the half-integral solution. Again, we rely on the integrality of matroid intersection polytopes to construct an approximately good integral solution from the given half-integral solution.

All together, we obtain an integral solution to the original instance of cost at most $e^{O(p)} \cdot z^*$.

A.1 Obtaining a half-integral solution

Step I: Consolidating Clients. In this section, we analyze the “client consolidation” subroutine, Algorithm 4, which outputs a new set of demands that is supported on a well-separated set of clients.

An important notion in the framework of [Charikar et al., 2002] is the *fractional distance of a client v to its facility* w.r.t. an optimal fractional solution (x, y) . In our setting with ℓ_p clustering cost, the fractional

distance of clients is defined as $\mathcal{R}(v) := (\sum_{u \in \mathcal{F}} d(v, u)^p \cdot x_{vu})^{1/p}$. In particular, if (x, y) is an integral solution then $\mathcal{R}(v)$ denotes the distance of client v to the facility it is assigned to. In other words, $\mathcal{R}(v)$ is the assignment cost of one unit of demand at client v .

We consider the clients v_1, \dots, v_n in a non-decreasing order of their fractional distances: $\mathcal{R}(v_1) \leq \dots \leq \mathcal{R}(v_n)$. At the time we are processing client v_i with *non-zero* demand, we check whether there exist another client v_j with non-zero demand such that $j > i$ and $d(v_i, v_j) \leq 2^{\frac{p+1}{p}} \cdot \mathcal{R}(v_j)$. If there exists such a client, then we add the demand of v_j to v_i and set the demand of v_j to zero.

Observe that throughout the client consolidation subroutine, once a client is processed, the algorithm never moves its demand in the rest of the procedure. This in particular implies that during the client consolidation subroutine, the demand of each client moves at most once.

Algorithm 4 consolidating clients.

```

1: Input:  $(x, y)$  is a feasible solution of FACILITYMATLP( $w, \mathcal{M}$ )
2:  $\mathcal{R}(v) = (\sum_{u \in \mathcal{F}} d(v, u)^p \cdot x_{vu})^{1/p}$  for all  $v \in P$ 
3:  $w'(v) = w(v)$  for all clients  $v \in P$ 
4: sort the points in  $P$  so that  $\mathcal{R}(v_1) \leq \mathcal{R}(v_2) \leq \dots \leq \mathcal{R}(v_n)$ 
5: for  $i = 1$  to  $n - 1$  do
6:   for  $j = i + 1$  to  $n$  do
7:     if  $d(v_i, v_j) \leq 2^{\frac{p+1}{p}} \mathcal{R}(v_j)$  and  $w'(v_i) > 0$  then
8:        $w'(v_i) = w'(v_i) + w'(v_j)$ 
9:        $w'(v_j) = 0$ 
10:    end if
11:  end for
12: end for

```

Next we define $P' \subseteq P$ as the support of w' . The following claim shows that the clients in P' are well-separated.

Claim A.1 For every pair of clients $u, v \in P'$, $d(u, v) > 2^{\frac{p+1}{p}} \max(\mathcal{R}(v), \mathcal{R}(u))$

Proof: Suppose there exists a pair of clients $v, u \in P'$ such that $d(v, u) \leq 2^{\frac{p+1}{p}} \cdot \max(\mathcal{R}(v), \mathcal{R}(u))$. Wlog, assume that $\mathcal{R}(v) \leq \mathcal{R}(u)$. However, since $d(v, u) \leq 2^{\frac{p+1}{p}} \cdot \mathcal{R}(u)$, at the iteration in Algorithm 4 that processes client u , the algorithm will move the demand of u to v . Hence, at the end of the algorithm $w'(u) = 0$ which is a contradiction. \square

Next, we show that a feasible solution of FACILITYMATLP for the original instance is a feasible solution of FACILITYMATLP for the constructed well-separated instance with the same or smaller cost.

Lemma A.2 Let (x, y) be a feasible solution of FACILITYMATLP(w, \mathcal{M}) with cost z . Then, (x, y) is a feasible solution of FACILITYMATLP(w', \mathcal{M}) with cost at most z .

Proof: Since the set of constraints in FACILITYMATLP(w', \mathcal{M}) is the same as the set of constraints in FACILITYMATLP(w, \mathcal{M}), (x, y) is a feasible solution of FACILITYMATLP(w', \mathcal{M}). Moreover, in Algorithm 4, a client u moves its demand to another client v with a lower assignment cost (i.e., $\mathcal{R}(v) \leq \mathcal{R}(u)$). Hence, the cost of solution (x, y) on FACILITYMATLP(w', \mathcal{M}) is at most z . \square

Theorem A.3 Let F' be an integral solution of the well-separated instance (w', \mathcal{M}) of cost at most z' . Then, F' is a solution of the original instance (w, \mathcal{M}) of cost at most $4 \cdot 16^{p-1} \cdot z^* + (\frac{8}{7})^{p-1} \cdot z'$ where z^* is the optimal cost of FACILITYMATLP(w, \mathcal{M}).

Proof: Since the set of constraints in FACILITYMATLP(w, \mathcal{M}) is the same as the set of constraints in FACILITYMATLP(w', \mathcal{M}), F' is a feasible solution of FACILITYMATLP(w, \mathcal{M}). For every client $v \in P$, we assume that Algorithm 4 has moved the demand of v to $v' \in P'$. More precisely, the algorithm may either move the demand of v to another client $v' = u$ or keep it at the same client $v' = v$. Moreover, in both

cases, $d(v, v') \leq 2^{\frac{p+1}{p}} \mathcal{R}(v)$. Hence,

$$\begin{aligned} d(v, F')^p &\leq 8^{p-1} \cdot d(v, v')^p + \left(\frac{8}{7}\right)^{p-1} \cdot d(v', F')^p && \triangleright \text{Observation 2.2 with } \lambda = 7 \\ &\leq 4 \cdot 16^{p-1} \cdot \mathcal{R}(v)^p + \left(\frac{8}{7}\right)^{p-1} \cdot \mathcal{R}(v')^p && \triangleright \text{since } d(v, v') \leq 2^{\frac{p+1}{p}} \mathcal{R}(v) \end{aligned} \quad (10)$$

Thus, the cost of solution F' over the original instance is,

$$\begin{aligned} \sum_{u \in F'} f(u) + \sum_{v \in P} d(v, F')^p &\leq \sum_{u \in F'} f(u) + \sum_{v \in P} 4 \cdot 16^{p-1} \cdot \mathcal{R}(v)^p + \left(\frac{8}{7}\right)^{p-1} \cdot \mathcal{R}(v')^p && \triangleright \text{by Eq. (10)} \\ &\leq 4 \cdot 16^{p-1} \cdot z^* + \left(\frac{8}{7}\right)^{p-1} \cdot z' && \square \end{aligned}$$

Step II: Transforming to a half-integral solution. In this section, we provide a method to construct a half-integral solution (w', \mathcal{M}) from the optimal fractional solution (x, y) , previously denoted as (x^*, y^*) . More precisely, we start with the optimal fractional solution (x, y) and after two steps, construct a feasible half-integral solution of the well-separated instance whose cost is not more than $3^p \cdot z^*$.

First, we define a few useful notions for our algorithm in this section and its analysis. For every client $v \in P'$, we define $F(v)$ to be the set of all facilities u such that v is the closest client to u in P' , i.e. $F(v) := \{u \in \mathcal{F} : d(v, u) = \min_{s \in P'} d(s, u)\}$ with ties broken arbitrarily. Furthermore, let $F'(v) \subseteq F(v) := \{u \in F(v) : d(u, v) \leq 2^{1/p} \cdot \mathcal{R}(v)\}$. Lastly, for each client v , we define $\gamma_v := \min_{u \notin F(v)} d(v, u)$ and let $G(v) := \{u \in F(v) : d(v, u) \leq \gamma_v\}$.

First, we show that for every client v , all ‘‘nearby’’ facilities are contained in $F(v)$.

Lemma A.4 *For every client $v \in P'$, the set $F(v)$ contains all the facilities u such that $d(v, u) \leq 2^{1/p} \cdot \mathcal{R}(v)$.*

Proof: Suppose that there exists a facility $u \notin F(v)$ with $d(v, u) \leq 2^{1/p} \cdot \mathcal{R}(v)$. Hence, there exists a client $v' \in P'$ such that $d(v', u) \leq d(v, u) \leq 2^{1/p} \cdot \mathcal{R}(v)$. By the approximate triangle inequality for $d(\cdot, \cdot)^p$,

$$\begin{aligned} d(v, v')^p &\leq 2^{p-1} \cdot (d(v, u)^p + d(u, v')^p) \\ &\leq 2^p \cdot d(v, u)^p && \triangleright \text{since } d(v', u) \leq d(v, u) \\ &\leq 2^{p+1} \cdot \mathcal{R}(v)^p, \end{aligned}$$

which contradicts the well-separatedness property of clients in P' . \square

Corollary A.5 *For every client $v \in P'$, $\gamma_v > 2^{1/p} \cdot \mathcal{R}(v)$. In particular, for every client $v \in P'$, $F'(v) \subseteq G(v)$.*

Proof: Let $u \in \mathcal{F} \setminus F(v)$ be the facility such that $\gamma_v = d(u, v) \leq 2^{1/p} \cdot \mathcal{R}(v)$. Then, by Lemma A.4, facility u is element of $F(v)$, giving a contradiction. For the second part, consider a facility $u \in F'(v)$. Then, by definition of F' and the first part of this proof, $d(u, v) \leq 2^{1/p} \cdot \mathcal{R}(v) < \gamma_v$ and therefore $u \in G(v)$. \square

For convenience, we provide a simple subroutine that computes the optimal assignments of clients (i.e. x) for any given feasible (possibly *fractional*) set of open facilities y . Note that if y is integral (resp. half-integral), the resulting optimal assignment x is integral (resp. half-integral) too.

Next, we show another useful property of $F(\cdot)$ which is crucial in constructing the half-integral solution of the well-separated instance.

Claim A.6 *For every client $v \in P'$, $\sum_{u \in F'(v)} x_{vu} \geq 1/2$.*

Proof: To prove the statement, we show that $\sum_{u \notin F'(v)} x_{vu} \leq 1/2$.

$$\begin{aligned} \sum_{u \notin F'(v)} x_{vu} \cdot 2 \cdot \mathcal{R}(v)^p &\leq \sum_{u \notin F'(v)} x_{vu} \cdot d(u, v)^p && \triangleright \text{by Corollary A.5} \\ &\leq \sum_{u \in \mathcal{F}} x_{vu} \cdot d(u, v)^p = \mathcal{R}(v)^p \end{aligned}$$

Algorithm 5 constructs an optimal assignment of clients to a given set of facilities.

-
- 1: **Input:** open facilities y and demand function $w : P \rightarrow \mathbb{R}$
 - 2: **initialize** $x_{uv} \leftarrow 0$ **for all** $u \in \mathcal{F}, v \in P$.
 - 3: **while** there exists $v \in P$ with $w(v) > 0$ and $\sum_{u \in \mathcal{F}} x_{uv} < 1$ **do**
 - 4: **find** a closest facility u to v that satisfies $y_u - x_{uv} > 0$.
 - 5: $x_{uv} \leftarrow \min(1 - \sum_{u \in \mathcal{F}} x_{uv}, y_u)$.
 - 6: **end while**
 - 7: **return** x
-

Hence, $\sum_{u \notin F'(v)} x_{vu} \leq 1/2$. Since in an optimal solution of $\text{FACILITYMATLP}(w, \mathcal{M})$, for every client v , $\sum_{u \in \mathcal{F}} x_{vu} = 1$. Thus, we have that $\sum_{u \in F'(v)} x_{vu} \geq 1/2$. \square

Next, we show that in the well-separated instance, for every client $v \in P'$ there exists another client $v' \in P'$ whose “nearby” facilities are close to v as well. More precisely,

Claim A.7 *Consider a client $v \in P'$. Let $u \in F(v')$ be the facility such that $\gamma_v = d(v, u)$. Then, for every facility $u' \in F'(v')$, $d(v, u') \leq 3\gamma_v$.*

Proof: See Figure 1 for an illustration of the relevant distances. First, we bound $d(v, v')$.

$$\begin{aligned}
 d(v, v')^p &\leq 2^{p-1} \cdot (d(v, u)^p + d(u, v')^p) && \triangleright \text{approximate triangle inequality} \\
 &\leq 2^p \cdot d(v, u)^p && \triangleright \text{since } u \in F(v') \\
 &\leq 2^p \cdot \gamma_v^p && (11)
 \end{aligned}$$

Furthermore, since clients in P' are well-separated, by Claim A.1, $d(v, v')^p > 2^{p+1} \cdot \max(\mathcal{R}(v), \mathcal{R}(v'))^p$. Moreover, since $u' \in F'(v')$, by Lemma A.4, $d(v', u')^p \leq 2 \cdot \mathcal{R}(v')^p$. Hence,

$$\begin{aligned}
 d(v', u')^p &\leq 2 \cdot \mathcal{R}(v')^p && \triangleright \text{by Lemma A.4} \\
 &\leq 2 \cdot \max(\mathcal{R}(v), \mathcal{R}(v'))^p \\
 &\leq \frac{1}{2^p} \cdot d(v, v')^p && \triangleright \text{by the well-separateness property of } P' \text{ (Claim A.1)} \\
 &\leq \gamma_v^p && \triangleright \text{Eq. (11)}
 \end{aligned} \tag{12}$$

By an application of the general form of approximate triangle inequality for $d(\cdot, \cdot)^p$,

$$\begin{aligned}
 d(v, u')^p &\leq \left(\frac{3}{2}\right)^{p-1} \cdot d(v, v')^p + 3^{p-1} \cdot d(v', u')^p && \triangleright \text{Observation 2.2 with } \lambda = 2 \\
 &\leq 2 \cdot 3^{p-1} \cdot \gamma_v^p + 3^{p-1} \cdot \gamma_v^p && \triangleright \text{Eq. (11) and (12)} \\
 &\leq 3^p \cdot \gamma_v^p && \square
 \end{aligned}$$

Next, we prove the main theorem of this section.

Theorem A.8 *Let z^* denote the cost of an optimal solution of $\text{FACILITYMATLP}(w, \mathcal{M})$. There exists a half-integral solution of $\text{FACILITYMATLP}(w', \mathcal{M})$ of cost at most $3^p \cdot z^*$.*

We begin with an optimal solution (x, y) of $\text{FACILITYMATLP}(w, \mathcal{M})$ of cost z^* which by Lemma A.2 is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$ of cost at most z^* . The first step in the proof is to construct an “intermediate” feasible solution (x', y') as follows.

$$y'_u := \begin{cases} x_{vu} & \text{if there exists a client } v \in P' \text{ such that } u \in G(v) \\ 0 & \text{otherwise} \end{cases}$$

Note that if there is no $v \in P'$ such that $u \in G(v)$, then $y_u = 0$.

Claim A.9 *For every client $v \in P'$, $\sum_{u \in G(v)} y'_u \leq 1$.*

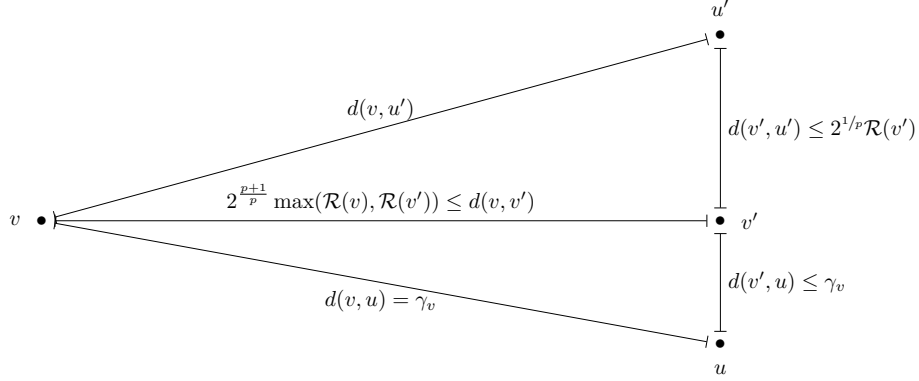


Figure 1: Illustration of the relevant distances to bound the distance between v and u' .

Proof: Note that since the sets of facilities $\{F(v)\}_{v \in P'}$ are disjoint and for each $v \in P'$, $G(v) \subseteq F(v)$, the sets $\{G(v)\}_{v \in P'}$ are disjoint too. Thus, by the way we constructed the solution y' ,

$$\sum_{u \in G(v)} y'_v = \sum_{u \in G(v)} x_{vu} \leq \sum_{u \in \mathcal{F}} x_{vu} = 1,$$

where the first equality follows from the disjointness of $\{G(v)\}_{v \in P'}$ and the second equality follows from the optimality of the solution (x, y) . \square

Next, following the approach of [Swamy, 2016], we introduce a *modified cost function* T that serves as a proxy for the actual cost:

$$T(y') = \sum_{u \in \mathcal{F}} y'_u \cdot f(u) + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot y'_u + 3^p \gamma^p (1 - \sum_{u \in G(v)} y'_u) \right) \quad (13)$$

This is crucial in bounding the cost of (x', y') , where x' is the optimal assignment w.r.t. y' and w' as constructed by Algorithm 5. Furthermore, the cost function T plays an important role in showing the existence of a good approximate half-integral solution of $\text{FACILITYMATLP}(w', \mathcal{M})$.

Lemma A.10 *The solution (x', y') is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$ and $\text{cost}(x', y') \leq T(y') \leq 3^p \cdot z^*$.*

Proof: For every facility $u \in \bigcup_{v \in P'} G(v)$, let v denote the client such that $u \in G(v)$. Note that the disjointness of $\{G(v)\}_{v \in P'}$ implies the uniqueness of such client. Note that for every $S \subseteq \mathcal{F}$,

$$\begin{aligned} \sum_{u \in S} y'_u &= \sum_{u \in \bigcup_{v \in P'} G(v) \cap S} y'_u + \sum_{u \in S \setminus \bigcup_{v \in P'} G(v)} y'_u \\ &= \sum_{u \in \bigcup_{v \in P'} G(v) \cap S} x_{vu} && \triangleright \text{by the definition of } y' \\ &\leq \sum_{u \in S} y_u \leq r(S), \end{aligned}$$

where the last two inequalities hold since (x, y) is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$. Hence, (x', y') satisfies the matroid constraint \mathcal{M} (i.e., constraint (8) in FACILITYMATLP).

Furthermore, line 5 of Algorithm 5 ensures the constraints (7) and (9) of $\text{FACILITYMATLP}(w', \mathcal{M})$ are satisfied by (x', y') . Hence, (x', y') is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$.

To prove $\text{cost}(x', y') \leq T(y')$ we show that y' is contained in the polytope \mathcal{P} —the polytope \mathcal{P} is defined formally in (16). Therefore, by Lemma A.12, $T(y') \geq \text{cost}(x', y')$. The first condition of \mathcal{P} encodes the

matroid independence constraint with $\sum_{u \in S} y'_u \leq r(S)$. As already stated in the context of the feasibility of (x', y') for $\text{FACILITYMATLP}(w', \mathcal{M})$, y' satisfies the matroid constraint. Secondly, for each client $v \in P'$,

$$\begin{aligned} \sum_{u \in F'(v)} y'_u &= \sum_{u \in F'(v)} x_{vu} && \triangleright \text{by the definition of } y' \text{ and since } \forall v \in P', F'(v) \subseteq G(v) \\ &\geq 1/2 && \triangleright \text{by Claim A.6} \end{aligned}$$

Finally, by Claim A.9, for every $v \in P'$, $\sum_{u \in G(v)} y'_u \leq 1$. Hence, $y' \in \mathcal{P}$.

Next, we show that $T(y') \leq 3^p \cdot z^*$.

$$\begin{aligned} T(y') &= \sum_{u \in \mathcal{F}} f(u) \cdot y'_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot y'_u + 3^p \gamma_v^p (1 - \sum_{u \in G(v)} y'_u) \right) \\ &\leq \sum_{u \in \mathcal{F}} f(u) \cdot y'_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot x_{vu} + 3^p \sum_{u \in \mathcal{F} \setminus G(v)} d(v, u)^p \cdot x_{vu} \right) \end{aligned} \quad (14)$$

$$\begin{aligned} &\leq \sum_{u \in \mathcal{F}} f(u) \cdot y_u + 3^p \sum_{v \in P'} w'(v) \sum_{u \in \mathcal{F}} d(v, u)^p \cdot x_{vu} \\ &\leq 3^p \cdot z^* \end{aligned} \quad (15) \quad \triangleright \text{by Lemma A.2}$$

Inequality (14) holds since by the definition of $G(v)$, for every client $v \in P'$ and facility $u \in \mathcal{F} \setminus G(v)$, $d(v, u) > \gamma_v$. Inequality (15) holds since for any non-zero y'_u there exist a client v such that $y'_u \leq x_{vu} \leq y_u$ —note that if $y'_u = 0$ then $y'_u \leq y_u$ holds trivially. \square

Besides the fact that the modified cost function T provides an upper bound for the actual cost of *certain solutions* of the well-separated instance, by the standard results for the matroid intersection problem, we can find a half-integral solution (x'', y'') whose modified cost T is minimized. In particular, we can find a half-integral feasible solution (x'', y'') of the well-separated instance such that $T(y'') \leq T(y')$. Before describing our method for constructing (x'', y'') , we formally define the set of solutions from which we choose the half-integral solution (x'', y'') .

$$\mathcal{P} := \left\{ y \in \mathbb{R}_+^{\mathcal{F}} : \sum_{u \in S} y_u \leq r(S) \quad \forall S \subseteq \mathcal{F}, \quad 1/2 \leq \sum_{u \in F'(v)} y_u, \quad \sum_{u \in G(v)} y_u \leq 1 \quad \forall v \in P' \right\} \quad (16)$$

Note that $y' \in \mathcal{P}$ (it is formally proved in the proof of Lemma A.10). First, we show that for any solution $\bar{y} \in \mathcal{P}$ and its optimal assignment \bar{x} w.r.t. \bar{y} and w' (e.g. as described in Algorithm 5), $\text{cost}(\bar{x}, \bar{y}) \leq T(\bar{y})$.

Claim A.11 *For every feasible solution $\bar{y} \in \mathcal{P}$ and any feasible assignment \tilde{x} w.r.t. \bar{y} and w' , the solution (\tilde{x}, \bar{y}) is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$.*

Proof: Since $\bar{y} \in \mathcal{P}$, it trivially satisfies the matroid constraint \mathcal{M} (i.e., constraint (8) in FACILITYMATLP). Furthermore, given that \tilde{x} is a feasible assignment w.r.t. \bar{y} and w' , (\tilde{x}, \bar{y}) satisfies constraints (7) and (9) of $\text{FACILITYMATLP}(w', \mathcal{M})$. Hence, (\tilde{x}, \bar{y}) is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$. \square

Lemma A.12 *For every solution $\bar{y} \in \mathcal{P}$ and its optimal assignment \bar{x} w.r.t. w' , $\text{cost}(\bar{x}, \bar{y}) \leq T(\bar{y})$.*

Proof: For every client $v \in P'$ let v' denote the client guaranteed by Claim A.7; $\forall u' \in F'(v')$, $d(v, u') \leq 3\gamma_v$. Moreover, we construct an assignment of \bar{y} denoted as \hat{x} as follows. For each $v \in P'$, $\hat{x}_{vu} := \bar{x}_{vu}$ if $u \in G(v)$. Next, we consider the facilities in $F'(v')$ in an arbitrary order u'_1, \dots, u'_ℓ and process them in this order one by one. For each $j \leq \ell$, we set $\hat{x}_{vu'_j} := \min(y'_{u'_j}, (1 - \sum_{u \in G(v)} \hat{x}_{vu} - \sum_{i < j} \hat{x}_{vu'_i}))$. Finally, for the remaining facilities $u' \in \mathcal{F} \setminus (G(v) \cup F'(v'))$, we set $\hat{x}_{vu'} = 0$. Since for each client $v \in P'$, $1/2 \leq \sum_{u \in F'(v)} y'_u \leq \sum_{u \in G(v)} y'_u$, the constructed assignment \hat{x} is a feasible assignment—i.e., (\hat{x}, \bar{y}) is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$. Moreover, our construction ensures that for every client $v \in P'$, $\sum_{u \in \mathcal{F}} \hat{x}_{vu} = \sum_{u \in G(v) \cup F'(v')} \hat{x}_{vu} = 1$. Finally, by the optimality of the assignment \bar{x} w.r.t. \bar{y} and w' , $\text{cost}(\bar{x}, \bar{y}) \leq \text{cost}(\hat{x}, \bar{y})$.

$$\begin{aligned}
 \text{cost}(\bar{x}, \bar{y}) &\leq \text{cost}(\hat{x}, \bar{y}) \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} w'(v) \cdot \sum_{u \in \mathcal{F}} d(v, u)^p \cdot \hat{x}_{vu} \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot \hat{x}_{vu} + \sum_{u \in F'(v')} d(v, u)^p \cdot \hat{x}_{vu} \right) \\
 &\leq \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot \hat{x}_{vu} + 3^p \gamma_v^p \sum_{u \in F'(v')} \hat{x}_{vu} \right) \quad \triangleright \text{by Claim A.7} \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot \hat{x}_{vu} + 3^p \gamma^p \left(1 - \sum_{u \in G(v)} \hat{x}_{vu}\right) \right) \quad \triangleright \sum_{u \in G(v) \cup F'(v')} \hat{x}_{vu} = 1 \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in G(v)} d(v, u)^p \cdot \bar{x}_{vu} + 3^p \gamma^p \left(1 - \sum_{u \in G(v)} \bar{x}_{vu}\right) \right) \quad \triangleright \forall u \in G(v), \hat{x}_{vu} = \bar{x}_{vu} \\
 &= T(\bar{y}) \quad \square
 \end{aligned}$$

Next, we show that there exist a half-integral solution y'' that minimizes the modified cost function T over the set of solutions described by \mathcal{P} .

Lemma A.13 *There is a half-integral solution y'' that minimizes T over the polytope \mathcal{P} .*

Proof: For the proof of this Lemma, we refer to [Swamy, 2016, Appendix A] where it is shown that the polytope \mathcal{P} has half-integral extreme solution. Hence, there is a polynomial time algorithm to find a half-integral solution $y'' \in \mathcal{P}$ that minimizes the *linear* cost function T .

Note that by Claim A.11, (x'', y'') is a feasible solution of $\text{FACILITYMATLP}(w', \mathcal{M})$ where x'' is the optimal (feasible) assignment w.r.t. y'' and w' . \square

Finally, we have all the pieces to prove Theorem A.8.

Proof of Theorem A.8: By Lemma A.10, (x', y') is a feasible solution to $\text{FACILITYMATLP}(w', \mathcal{M})$ and its cost is at most $T(y') \leq 3^p \cdot z^*$. Moreover, Lemma A.10 shows that the solution y' is contained in the polytope \mathcal{P} . Then, by an application of Lemma A.13, there exists a half-integral solution y'' such that $T(y'') \leq T(y')$ —in fact, the solution y'' minimizes T in the polytope \mathcal{P} . Now, we consider the half-integral solution (x'', y'') of $\text{FACILITYMATLP}(w', \mathcal{M})$ where x'' is the optimal assignment w.r.t. y'' and w' .

$$\begin{aligned}
 \text{cost}(x'', y'') &\leq T(y'') \quad \triangleright \text{by Lemma A.12} \\
 &\leq T(y') \quad \triangleright \text{by the optimality of } y'' \text{ w.r.t. } T \text{ in the polytope } \mathcal{P} \\
 &\leq 3^p \cdot z^* \quad \triangleright \text{by Lemma A.10} \quad \square
 \end{aligned}$$

A.2 Converting (x'', y'') to an integer solution

In this section, we show how to convert the half-integral solution (x'', y'') of the well-separated instance to an integral solution of the well-separated instance without losing more than $e^{O(p)} \cdot z^*$ in the cost.

Theorem A.14 *Let z^* denote the cost of an optimal solution of $\text{FACILITYMATLP}(w, \mathcal{M})$. There exists an integral solution of $\text{FACILITYMATLP}(w', \mathcal{M})$ of cost at most $(4 \cdot 3^{p-1} + 2) \cdot 3^p \cdot z^*$.*

Similarly to the notion *fractional distance* \mathcal{R} defined w.r.t. the optimal solution of (x, y) of the original instance (see Step I in Section A.1), for every client $v \in P'$, we define the fractional distance of v w.r.t. (x'', y'') as $\mathcal{R}''(v) := \left(\sum_{u \in \mathcal{F}} d(u, v)^p \cdot x''_{uv} \right)^{\frac{1}{p}}$. Moreover, for each client $v \in P'$, we denote the set of *serving* facilities of v in (x'', y'') as $\mathcal{F}''(v) := \{u \in \mathcal{F} : x''_{vu} > 0\}$.

Step III: Identify core clients. First, in Algorithm 6, we construct a subset of clients $P'' \subseteq P'$, called *core clients* and a mapping $cr : P' \rightarrow P''$. The crucial property of the core clients P'' is the following: every

Algorithm 6 constructs a set of core clients.

```

1: Input:  $(x'', y'')$ : Half-integral solution of FACILITYMATLP( $w', \mathcal{M}$ ) from Theorem A.8
2:  $\mathcal{R}''(v) \leftarrow (\sum_{u \in \mathcal{F}} d(v, u)^p \cdot x''_{vu})^{1/p}$  for all  $v \in P'$ 
3:  $\mathcal{F}''(v) \leftarrow \{u : x''_{uv} > 0\}$  for all clients  $v \in P'$ 
4:  $P'' \leftarrow \emptyset$ 
5: while  $P' \neq \emptyset$  do
6:    $v^* = \operatorname{argmin}_{v \in P'} \mathcal{R}''(v)$ 
7:    $P' \leftarrow P' \setminus \{v^*\}$ ,  $P'' \leftarrow P'' \cup \{v^*\}$ ,  $cr(v^*) \leftarrow v^*$ 
8:   for all  $v' \in P'$  do
9:     if  $\mathcal{F}''(v^*) \cap \mathcal{F}''(v') \neq \emptyset$  then
10:        $P' \leftarrow P' \setminus \{v'\}$ ,  $cr(v') = v^*$ 
11:     end if
12:   end for
13: end while
14: return  $P''$ ,  $cr$ 
    
```

facility $u \in \mathcal{F}$ is serving *at most* one client in P'' . In other words the family of sets $\{\mathcal{F}''(v)\}_{v \in P''}$ are disjoint.

Claim A.15 For every client $v \in P'$, $\mathcal{R}''(cr(v)) \leq \mathcal{R}''(v)$.

Proof: The inequality holds trivially for the core clients $v \in P''$. Let $v' \in P' \setminus P''$. At the iteration in which $cr(v)$ is added to P'' , v is still present in P' . Hence, by the condition at line 6, $\mathcal{R}''(cr(v)) \leq \mathcal{R}''(v)$. \square

Step IV: Obtaining an integral solution (\tilde{x}, \tilde{y}) . Similarly to our approach for constructing the half-integral solution (x'', y'') , we first construct an “intermediate” solution \tilde{y}' with certain structures. Later, we exploit the known results in matroid intersection to find a “good” integral solution for the constructed intermediate solution.

$$\tilde{y}'_u := \begin{cases} x''_{vu} & \text{if there exists a client } v \in P'' \text{ such that } u \in \mathcal{F}''(v) \\ y''_u & \text{otherwise} \end{cases}$$

Lemma A.16 The solution (\tilde{x}', \tilde{y}') where \tilde{x}' is an optimal assignment w.r.t. \tilde{y}' and w' is a feasible solution of FACILITYMATLP(w', \mathcal{M}).

Proof: The constraints (7) and (9) are satisfied by the way \tilde{x}' is constructed via Algorithm 5. The matroid constraint, constraint (8), holds since for every facility $u \in \mathcal{F}$, $\tilde{y}'_u \leq y''_u$ and y'' satisfies the matroid constraint. \square

Claim A.17 For every core client $v \in P''$, $\sum_{u \in \mathcal{F}''(v)} \tilde{y}'_u = 1$.

Proof:

$$\begin{aligned} \sum_{u \in \mathcal{F}''(v)} \tilde{y}'_u &= \sum_{u \in \mathcal{F}''(v)} x''_{vu} && \triangleright \text{by the definition of } \tilde{y}' \\ &= \sum_{u \in \mathcal{F}} x''_{vu} && \triangleright \mathcal{F}''(v) := \{u \in \mathcal{F} : x''_{vu} > 0\} \\ &= 1 && \triangleright \text{by the feasibility of } (x'', y'') \text{ for FACILITYMATLP} \quad \square \end{aligned}$$

Next, for every client $v \in P'$, we define the *primary* p_v and the *secondary* s_v facilities of v such that p_v denotes the nearest facilities to v among the possibly two facilities serving v in the half-integral solution (x'', y'') . Note that since (x'', y'') is a half-integral solution, either $x''_{vp_v} = x''_{vs_v} = \frac{1}{2}$ or $x''_{vp_v} = 1$ otherwise. For technical reason, in the latter case, we set $s_v = p_v$.

Claim A.18 For every client $v \in P'$, $\mathcal{R}''(v)^p = \frac{1}{2}(d(p_v, v)^p + d(s_v, v)^p)$.

Proof: It simply follows from the half-integrality of the solution (x'', y'') and the definition of the primary and the secondary facilities. \square

Claim A.19 For every client $v \in P'$, $d(p_v, v)^p \leq \mathcal{R}''(v)^p \leq d(s_v, v)^p \leq 2\mathcal{R}''(v)^p$.

Proof: By the definition of p_v and s_v , $d(p_v, v)^p \leq \frac{1}{2}(d(p_v, v)^p + d(s_v, v)^p) \leq d(s_v, v)^p$. By an application of Claim A.18 the proof is complete. \square

Similarly to the previous section, we introduce a cost function H that serves as a proxy to bound the cost of a set of half-integral solutions we are considering in this section. Let $H(\tilde{y}') := \sum_{u \in \mathcal{F}} f(u) \cdot \tilde{y}'_u + \sum_{v \in P'} w'(v) A_v(\tilde{y}')$ where $A_v(\tilde{y}')$ is the proxy for the per-unit assignment cost of a client $v \in P'$ which is defined as

$$A_v(\tilde{y}') := \sum_{u \in \mathcal{F}''(cr(v))} d(u, v)^p \cdot \tilde{y}'_u \quad (17)$$

Lemma A.20 The proxy cost of the intermediate solution \tilde{y}' is at most $(4 \cdot 3^{p-1} + 2)$ times the cost of (x'', y'') ; $H(\tilde{y}') \leq (4 \cdot 3^{p-1} + 2) \cdot \text{cost}(x'', y'')$.

Proof: Since for every facility $u \in \mathcal{F}$, $\tilde{y}'_u \leq y''_u$, $\sum_{u \in \mathcal{F}} f(u) \cdot \tilde{y}'_u \leq \sum_{u \in \mathcal{F}} f(u) \cdot y''_u$. Next, we consider the following cases to bound the contribution of the assignment cost of a client v in $H(\tilde{y}')$.

1. v is a core client ($v \in P''$)

$$\begin{aligned} A_v(\tilde{y}') &= \sum_{u \in \mathcal{F}''(cr(v))} d(v, u)^p \cdot \tilde{y}'_u \\ &= \sum_{u \in \mathcal{F}''(v)} d(v, u)^p \cdot \tilde{y}'_u && \triangleright cr(v) = v \\ &= \sum_{u \in \mathcal{F}''(v)} d(v, u)^p \cdot x''_{vu} && \triangleright \text{by the definition of } \tilde{y}' \\ &= \sum_{u \in \mathcal{F}} d(v, u)^p \cdot x''_{vu} && \triangleright \mathcal{F}''(v) := \{u \in \mathcal{F} : x''_{vu} > 0\} \\ &= \mathcal{R}''(v)^p \end{aligned}$$

2. v is not a core client ($v \in P \setminus P''$) and $p_v \in \mathcal{F}''(cr(v))$. Let $u^* \in \mathcal{F}''(cr(v)) \setminus \{p_v\}$.

$$\begin{aligned} A_v(\tilde{y}') &= \sum_{u \in \mathcal{F}''(cr(v))} d(v, u)^p \cdot \tilde{y}'_u \\ &= d(v, p_v)^p \cdot \tilde{y}'_{p_v} + d(v, u^*)^p \cdot \tilde{y}'_{u^*} && \triangleright \mathcal{F}''(cr(v)) = \{u^*, p_v\} \\ &\leq d(v, p_v)^p + 3^{p-1} \cdot (d(v, p_v)^p + d(p_v, cr(v))^p + d(cr(v), u^*)^p) && \triangleright \text{Eq. (2) and } \|\tilde{y}'\|_\infty \leq 1 \\ &\leq \mathcal{R}''(v)^p + 3^{p-1} \cdot (\mathcal{R}''(v)^p + 2\mathcal{R}''(cr(v))^p) && \triangleright \text{Claim A.18 and A.19} \\ &\leq \mathcal{R}''(v)^p + 3^p \cdot \mathcal{R}''(v)^p && \triangleright \text{Claim A.15} \\ &\leq (3^p + 1) \cdot \mathcal{R}''(v)^p \end{aligned}$$

3. v is not a core client ($v \in P \setminus P''$) and $p_v \notin \mathcal{F}''(cr(v))$. Since $p_v \notin \mathcal{F}''(cr(v))$, we have that

$s_v \in p_v \notin \mathcal{F}''(cr(v))$. Let $u^* \in \mathcal{F}''(cr(v)) \setminus \{s_v\}$.

$$\begin{aligned}
 A_v(\tilde{y}') &= \sum_{u \in \mathcal{F}''(cr(v))} d(v, u)^p \cdot \tilde{y}'_u \\
 &= d(v, s_v)^p \cdot \tilde{y}'_{s_v} + d(v, u^*)^p \cdot \tilde{y}'_{u^*} &> \mathcal{F}''(cr(v)) = \{u^*, s_v\} \\
 &\leq 2\mathcal{R}''(v)^p + d(v, u^*)^p &> \text{Claim A.19, } \tilde{y}'_{s_v} \leq 1 \\
 &\leq 2\mathcal{R}''(v)^p + 3^{p-1} \cdot (d(v, s_v)^p + d(s_v, cr(v))^p + d(cr(v), u^*)^p) &> \text{Eq. (2)} \\
 &\leq 2\mathcal{R}''(v)^p + 3^{p-1} \cdot (2\mathcal{R}''(v)^p + 2\mathcal{R}''(cr(v)))^p &> \text{Claim A.18 and A.19} \\
 &\leq 2\mathcal{R}''(v)^p + 4 \cdot 3^{p-1} \cdot \mathcal{R}''(v)^p &> \text{Corollary A.15} \\
 &\leq (4 \cdot 3^{p-1} + 2) \cdot \mathcal{R}''(v)^p
 \end{aligned}$$

Hence, summing over all clients in P' ,

$$\begin{aligned}
 H(\tilde{y}') &= \sum_{u \in \mathcal{F}} f(u) \cdot \tilde{y}'_u + \sum_{v \in P'} w'(v) \cdot A_v(\tilde{y}') \\
 &\leq \sum_{u \in \mathcal{F}} f(u) \cdot \tilde{y}'_u + (4 \cdot 3^{p-1} + 2) \cdot \sum_{v \in P'} w'(v) \cdot \mathcal{R}''(v)^p \\
 &\leq \sum_{u \in \mathcal{F}} f(u) \cdot y''_u + (4 \cdot 3^{p-1} + 2) \cdot \sum_{v \in P'} w'(v) \cdot \left(\sum_{u \in \mathcal{F}} d(u, v)^p \cdot x''_{uv} \right) \\
 &\leq (4 \cdot 3^{p-1} + 2) \cdot \text{cost}(x'', y'') \quad \square
 \end{aligned}$$

Next, we define the following polytope \mathcal{Q} that has integral extreme points and contains \tilde{y}' .

$$\mathcal{Q} := \left\{ y \in \mathbb{R}_+^{\mathcal{F}} : \sum_{u \in S} y_u \leq r(S) \quad \forall S \subseteq \mathcal{F}, \quad \sum_{u \in \mathcal{F}''(v)} y_u = 1 \quad \forall v \in P'' \right\} \quad (18)$$

First we show that for every solution $\bar{y} \in \mathcal{Q}$, $H(\bar{y}) \geq \text{cost}(\bar{x}, \bar{y})$ where \bar{x} is an optimal assignment w.r.t. \bar{y} and w' as described in Algorithm 5. We now prove that the cost of any vector $\tilde{y} \in \mathcal{Q}$ with its optimal assignment \tilde{x} obtained by Algorithm 5 is at most $H(\tilde{y})$. This Lemma proves for both \tilde{y}' and for \tilde{y} that $H(\tilde{y}')$ and respectively $H(\tilde{y})$ is an upper bound on the assignment cost.

Lemma A.21 *For every $\bar{y} \in \mathcal{Q}$, $\text{cost}(\bar{x}, \bar{y}) \leq H(\bar{y})$ where \bar{x} is an optimal assignment w.r.t. \bar{y} and w' .*

Proof: The total contribution of the facility opening cost in $\text{cost}(\bar{x}, \bar{y})$ and $H(\bar{y})$ are the same. Observe that for every client $v' \in P'$, there exists a core client $v \in P''$ such that $cr(v') = v$. In the following, we construct a feasible assignment \hat{x} w.r.t. \bar{y} and P'' such that its assignment cost is equal to the assignment cost of $H(\tilde{y}')$ (which is equal to $\sum_{v \in P'} A_v(\bar{y})$). Once we have \hat{x} , the lemma simply follows from the optimality of assignment \bar{x} w.r.t. \bar{y} and P'' .

We construct \hat{x} as follows. For each client $v \in P'$, $\hat{x}_{vu} = \bar{y}_u$ if $u \in \mathcal{F}''(cr(v))$ and zero otherwise. This is a feasible assignment w.r.t. \bar{y} and w' because $\bar{y} \in \mathcal{Q}$. We next bound the cost of solution (\hat{x}, \bar{y}) .

$$\begin{aligned}
 \text{cost}(\bar{x}, \bar{y}) &\leq \text{cost}(\hat{x}, \bar{y}) &> \text{by the optimality of } \bar{x} \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} \sum_{u \in \mathcal{F}} w' \cdot d(v, u)^p \cdot \hat{x}_{vu} \\
 &= \sum_{u \in \mathcal{F}} f(u) \cdot \bar{y}_u + \sum_{v \in P'} \sum_{u \in \mathcal{F}''(cr(v))} w'(v) \cdot d(v, u)^p \cdot \bar{y}_u &> \text{by the definition of } \hat{x} \\
 &= H(\bar{y}) \quad \square
 \end{aligned}$$

Lemma A.22 *There is an integral solution \tilde{y} that minimizes H over the polytope \mathcal{Q} .*

Proof: The desired solution \tilde{y} exists since H is a linear function and the extreme point of the polytope \mathcal{Q} are integral. The latter holds since \mathcal{Q} is non-empty and an intersection of two matroid polytopes (defined by \mathcal{M} and the partition matroid corresponding to $\sum_{u \in \mathcal{F}''(v)} y_u = 1, \forall v \in P''$). \square

Proof of Theorem A.14 By Theorem A.8, there exist a half-integral solution (x'', y'') of $\text{FACILITYMATLP}(\mathcal{M}, w')$ of cost at most $3^p \cdot z^*$. Let \tilde{y} be the minimizer of H over the polytope \mathcal{Q} . Moreover, let \tilde{x} denote the optimal assignment of \tilde{y} w.r.t. w' . By Lemma A.22, the solution (\tilde{x}, \tilde{y}) is an integral feasible solution of $\text{FACILITYMATLP}(\mathcal{M}, w')$. Furthermore,

$$\begin{aligned}
 \text{cost}(\tilde{x}, \tilde{y}) &\leq H(\tilde{y}) && \triangleright \text{by Lemma A.21} \\
 &\leq H(\tilde{y}') && \triangleright \text{since } \tilde{y} = \operatorname{argmin}_{y \in \mathcal{Q}} H(y) \text{ and } \tilde{y}' \in \mathcal{Q} \\
 &\leq (4 \cdot 3^{p-1} + 2) \cdot \text{cost}(x'', y'') && \triangleright \text{by Lemma A.20} \\
 &\leq (4 \cdot 3^{p-1} + 2) \cdot 3^p \cdot z^* && \triangleright \text{by Theorem A.8} \quad \square
 \end{aligned}$$

Now we are ready to state the main theorem of ℓ_p -norm facility location under matroid constraint.

Theorem A.23 (Main Theorem of ℓ_p -norm Facility Location Under Matroid Constraint) *For $p > 1$, there exists a polynomial time algorithm that finds a (16^p) -approximate solution of ℓ_p -clustering on (w, P) under matroid constraint \mathcal{M} .*

Proof: Following the result of this section, we first construct a well-separated instance (w', P') . By Theorem A.14, we can construct an integral solution of the well-separated instance of cost at most $(4 \cdot 3^{p-1} + 2) \cdot 3^p \cdot z^*$. Next, by Theorem A.3, the constructed solution can be extended to a feasible solution of the original instance (w, P) of cost at most $4 \cdot 16^{p-1} \cdot z^* + (\frac{8}{7})^{p-1} \cdot (4 \cdot 3^{p-1} + 2) \cdot 3^p \cdot z^* < 16^p \cdot z^*$ (for $p > 1$). \square

Remark A.24 Note that our approach works for $p = 1$ too and achieves a 22-approximation guarantee. However, since the result of [Swamy, 2016] provides an 8-approximation in this case ($p = 1$), we only consider $p > 1$ here.

B Missing Proofs

Lemma B.1 (Lemma A.1 Makarychev et al. [2019]) *Let x, y_1, \dots, y_n be non-negative real numbers and $\lambda > 0, p \geq 1$. Then,*

$$\left(x + \sum_{i=1}^n y_i\right)^p \leq (1 + \lambda)^{p-1} x^p + \left(\frac{(1 + \lambda)n}{\lambda}\right)^{p-1} \sum_{i=1}^n y_i^p.$$

Proof of Lemma 3.8. Let $u, v, w \in (\mathcal{F} \cup \mathcal{M})$ three arbitrary points and let u_P, v_P, w_P be their corresponding points from P . Furthermore, let $\hat{\varepsilon} := \min\{(\frac{\varepsilon(n-k)}{\beta \cdot k})^{1/p}, 1\}$.

First we prove that $d'(u, v) = 0 \iff u = v$. If $u = v$, then by line 13 the distance $d'(u, v)$ is set to zero. To show the other direction, if $d'(u, v) = 0$ then the constraint $u = v$ for the assignment in line 13 is satisfied since $d(u_p, v_p) > 0$ for all $u_p \neq v_p$ (line 14) and $d'(u, v) = \hat{\varepsilon} \cdot \delta > 0$ when $u_p = v_p$ and $u \neq v$ (line 15).

Secondly, we prove the symmetric property $d'(u, v) = d'(v, u)$. If $d'(u, v) = 0$, then by the first part $u = v$ and therefore $d'(v, u) = 0 = d'(u, v)$. Assume $d'(u, v) > 0$ which implies $u \neq v$. If $u_P \neq v_P$, then by line 14 and the metric properties of d , $d'(u, v) = d(u_P, v_P) = d(v_P, u_P) = d'(v, u)$ holds. Otherwise, by line 15, $d'(u, v) = \hat{\varepsilon} \cdot \delta = d'(v, u)$.

Lastly we show that the triangle inequality $d'(u, w) \leq d'(u, v) + d'(v, w)$ holds. If $u = w$ then by the first property, $d'(u, w) = 0$ so the inequality holds. Assume $u \neq w$ and consider their corresponding points u_P, w_P .

1. If $u_P = w_P$ then, $d'(u, w) = \hat{\varepsilon} \cdot \delta$. Let v_P be the corresponding point of v . If $v_P = u_P$, then $d'(u, v) = d'(u, w) = \hat{\varepsilon} \cdot \delta$ and therefore $d'(u, w) \leq d'(u, v) + d'(v, w)$ already holds. If $v_P \neq u_P$, then $d'(u, v) = d(u_P, v_P) \geq \min_{x, y \in P} d(x, y) \geq \hat{\varepsilon} \cdot \delta$. Thus $d'(u, w) \leq d'(u, v) + d'(v, w)$ holds.
2. If $u_P \neq w_P$ then $d'(u, w) = d(u_P, w_P) \geq \hat{\varepsilon} \cdot \delta$. Note that $(u_P = v_P \text{ and } v_P = w_P)$ can not hold, so consider the remaining three cases cases:

- (a) $v_P = w_P$ and $u_P \neq v_P$. Then $d'(u, w) = d'(u, v)$ and therefore $d'(u, w) \leq d'(u, v) + d'(v, w)$
- (b) $u_P = v_P$ and $v_P \neq w_P$. Then $d'(u, w) = d'(v, w)$ and therefore $d'(u, w) \leq d'(u, v) + d'(v, w)$
- (c) $u_P \neq v_P$ and $v_P \neq w_P$. Then $d'(u, w) = d(u_P, w_P)$, $d'(u, v) = d(u_P, v_P)$, $d'(v, w) = d(v_P, w_P)$ and since $d(\cdot)$ satisfies the triangle inequality, $d'(u, w) \leq d'(u, v) + d'(v, w)$ holds. \square

Proof of Theorem 3.11 Let CENTERPARTALG be a β -approximation algorithm for k -center under partition matroid constraint. Consider an instance of α -fair k -center on P and let $(P', \{(\overline{P}_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\})$ be the instance of k -center under partition matroid constraint constructed by Algorithm 3 with input parameters P, k and α . We show that the solution returned by CENTERPARTALG($P', \{(\overline{P}_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\}$) can be converted to a $(\beta + \varepsilon, 3)$ -bicriteria approximate solution of the given instance of α -fair k -center on P .

Let $\mathcal{B} = \{B_1, \dots, B_m\}$ be the critical regions of P constructed in Algorithm 1. Let SOL_C be the solution returned by CENTERPARTALG($P', \{(\overline{P}_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\}$) and let OPT be an optimal solution of α -fair k -center of P . Note that since adding centers in SOL_C only reduces the k -center cost of the solution on $(P', \{(\overline{P}_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\})$, without loss of generality we can assume that SOL_C picks exactly one center from each of B_i , for $i \in [m]$, and exactly $k - m$ centers from \overline{P}_0 . Now we construct a solution SOL of α -fair k -center on P using the solution SOL_C . We start with an initially empty set of centers SOL. In the first step, for each $B \in \mathcal{B}$, let c_i denote the center in $\text{SOL}_C \cap \overline{B}_i$ and then we add the point $c \in P$ corresponding to c_i to SOL. Next, in the second step, for each $o_0 \in \text{SOL}_C \cap \overline{P}_0$, we add the point $o \in P$ corresponding to o_0 to SOL. Note that as some of these points may have already been added to SOL in the first step, the final solution has at most k distinct centers.

Fairness approximation. By the first step in the construction of SOL, for each $i \in [m]$, $|B_i \cap \text{SOL}| \geq 1$. Hence, by Lemma 3.7, SOL is a (3α) -fair k -center clustering of P .

Cost approximation. First we show that the cost of SOL_C on P' is not smaller than the k -center clustering cost of P using SOL. Lets assume that there exist $v \in P$ such that $d(v, \text{SOL}) > d'(v', \text{SOL}_C)$ where v' is a copy of v in P' . Let c' be the closest center to v' in SOL_C . Let c denote the point in P corresponding to c' . Since after the second step of constructing SOL all original copies of the centers in SOL_C are added to SOL, $c \in \text{SOL}$. Hence, $d(v, \text{SOL}) \leq d(v, c) \leq d'(v', c') = d'(v', \text{SOL}_C)$ which is a contradiction. Hence, the cost of SOL_C is not smaller than the cost of SOL.

Next, we bound the cost of SOL_C on P' in terms of the cost of k -center clustering of P using OPT. By the definition of α -fairness, each point $v \in P$ must have a center in OPT within distance at most $\alpha \cdot r(v)$. Hence, for each critical region $B \in \mathcal{B}$, $|\text{OPT} \cap B| \geq 1$. For each $i \in [m]$, let c_i be the copy of an arbitrary center $c \in \text{OPT} \cap B_i$ in the set \overline{B}_i . For the remaining points in OPT, we pick their corresponding copies in the set \overline{P}_0 . Let OPT_C denote the constructed solution for the instance P' . Since OPT_C picks exactly one point from each set \overline{B}_i , for $i \in [m]$, and exactly $k - m$ points from \overline{P}_0 , OPT_C is a feasible solution for k -center under partition matroid constraint on instance $(P', \{(\overline{P}_0, k_0), (\overline{B}_1, k_1), \dots, (\overline{B}_m, k_m)\}, d')$. Moreover, since for every pair $(v \in P, c \in \text{OPT})$, there exists a pair $(v' \in P', c' \in \text{OPT}_C)$ such that $d'(v', c') \leq d(v, c) + \varepsilon \cdot \delta / \beta$, $\text{cost}_{k\text{center}}(\text{OPT}_C; P') \leq \text{cost}_{k\text{center}}(\text{OPT}; P) + \frac{\varepsilon \cdot \delta}{\beta}$. Hence,

$$\begin{aligned} \text{cost}_{k\text{center}}(\text{SOL}, P) &\leq \text{cost}_{k\text{center}}(\text{SOL}_C, P') \leq \beta \cdot \text{cost}_{k\text{center}}(\text{OPT}_C, P') \leq \beta \cdot (\text{cost}_{k\text{center}}(\text{OPT}, P) + \frac{\varepsilon \cdot \delta}{\beta}) \\ &\leq (\beta + \varepsilon) \cdot \text{cost}_{k\text{center}}(\text{OPT}, P), \end{aligned}$$

where the last inequality follows since $\text{cost}_{k\text{center}}(\text{OPT}, P) \geq \delta$. Thus, the k -center clustering cost of P using SOL is within a $\beta + \varepsilon$ factor of the cost of any optimal α -fair k -center of P . \square