# Ordered $k$-Median with Outliers

**Shichuan Deng** ✉
IIIS, Tsinghua University, Beijing, China

**Qianfan Zhang** ✉
IIIS, Tsinghua University, Beijing, China

## Abstract

We study a natural generalization of the celebrated ordered $k$-median problem, named *robust ordered $k$-median*, also known as ordered $k$-median with outliers. We are given facilities $\mathcal{F}$ and clients $\mathcal{C}$ in a metric space $(\mathcal{F} \cup \mathcal{C}, d)$, parameters $k, m \in \mathbb{Z}_+$ and a non-increasing non-negative vector $\boldsymbol{w} \in \mathbb{R}_+^m$. We seek to open $k$ facilities $F \subseteq \mathcal{F}$ and serve $m$ clients $C \subseteq \mathcal{C}$, inducing a *service cost vector* $\boldsymbol{c} = \{d(j, F) : j \in C\}$; the goal is to minimize the *ordered objective* $\boldsymbol{w}^\top \boldsymbol{c}^\downarrow$, where $d(j, F) = \min_{i \in F} d(j, i)$ is the minimum distance between client $j$ and facilities in $F$, and $\boldsymbol{c}^\downarrow \in \mathbb{R}_+^m$ is the non-increasingly sorted version of $\boldsymbol{c}$. Robust ordered $k$-median captures many interesting clustering problems recently studied in the literature, e.g., robust $k$-median, ordered $k$-median, etc.

We obtain the *first* polynomial-time constant-factor approximation algorithm for robust ordered $k$-median, achieving an approximation guarantee of 127. The main difficulty comes from the presence of outliers, which already causes an unbounded integrality gap in the natural LP relaxation for robust $k$-median. This appears to invalidate previous methods in approximating the highly non-linear ordered objective. To overcome this issue, we introduce a novel yet very simple reduction framework that enables linear analysis of the non-linear objective. We also devise the first constant-factor approximations for *ordered matroid median* and *ordered knapsack median* using the same framework, and the approximation factors are 19.8 and 41.6, respectively.

**2012 ACM Subject Classification** Theory of computation → Facility location and clustering

**Keywords and phrases** clustering, approximation algorithm, design and analysis of algorithms

**Digital Object Identifier** 10.4230/LIPIcs.APPROX/RANDOM.2022.34

**Category** APPROX

**Related Version** *Full Version*: https://arxiv.org/pdf/2011.04289.pdf

## 1 Introduction

$k$-supplier and $k$-median are two of the most fundamental clustering problems. In both problems, we are given facilities $\mathcal{F}$ and clients $\mathcal{C}$ in a metric space $(\mathcal{F} \cup \mathcal{C}, d)$ and a parameter $k \in \mathbb{Z}_+$; we need to select $k$ facilities $F \subseteq \mathcal{F}$, and only the objective functions are different. In $k$-supplier, the goal is to minimize the maximum distance from each client to its nearest facility in $F$, i.e., $\max_{j \in \mathcal{C}} d(j, F)$; $k$-supplier is NP-hard to approximate to a factor better than 3 [18], and a tight 3-approximation is given in [18]. In $k$-median, the objective is the sum of distances from each client to its nearest facility, i.e., $\sum_{j \in \mathcal{C}} d(j, F)$; $k$-median is NP-hard to approximate to a factor of $(1 + 2/e - \epsilon)$ for every $\epsilon > 0$ [20], and several constant-factor approximations are developed [2, 9, 11, 14, 21, 26]; currently the best approximation guarantee is $(2.675 + \epsilon)$ due to Byrka et al. [4].

Under the basic input $(\mathcal{F}, \mathcal{C}, d, k)$, let $\boldsymbol{c}_0 = \{d(j, F) : j \in \mathcal{C}\}$ be the service cost vector induced by solution $F$. The theoretical computer science community has lately shown increasingly more interests in clustering problems with more nuanced objective functions than $k$-supplier and $k$-median. For example, the ordered $k$-median problem (O$k$Med) naturally unifies these two problems via the ordered objective $\boldsymbol{w}_0^\top \boldsymbol{c}_0^\downarrow$, where $\boldsymbol{c}_0^\downarrow$ is the non-increasingly sorted version of $\boldsymbol{c}_0$ and $\boldsymbol{w}_0$ is a given non-increasing non-negative vector; it is easy to see that O$k$Med recovers $k$-supplier and $k$-median using only 0-1 vectors for $\boldsymbol{w}_0$. Several constant-factor approximations have been developed for O$k$Med [5, 7], and currently the best ratio is $(5+\epsilon)$ due to Chakrabarty and Swamy [8].

Meanwhile, a parallel line of research called robust clustering (also known as clustering with outliers) also attracts a lot of attention. These problems allow us to discard a certain number of clients and define the clustering objective on the remaining clients. In *robust $k$-center* (R$k$Cen), we are given an additional integer parameter $m \leq |\mathcal{C}|$ besides the basic input $(\mathcal{F}, \mathcal{C}, d, k)$ where $\mathcal{F} = \mathcal{C}$; we need to open $k$ facilities $F \subseteq \mathcal{F}$ and choose $m$ clients $C \subseteq \mathcal{C}$, and the objective is the maximum service cost in $C$, i.e., $\max_{j \in C} d(j, F)$. Charikar et al. [10] give a 3-approximation algorithm for R$k$Cen. Chakrabarty et al. [6] improve the result to a best-possible 2-approximation (also see Harris et al. [17]). It is easy to see that the objective of R$k$Cen is equivalent to $\boldsymbol{e}_{|\mathcal{C}|-m+1}^{\top} \boldsymbol{c}_0^{\downarrow}$, where $\boldsymbol{e}_t = \{0, \ldots, 0, 1, 0, \ldots, 0\}$ is the all-zero vector except for its $t$-th coordinate, which is 1. In *robust $k$-median* (R$k$Med), the input is the same as R$k$Cen except that $\mathcal{C}$ and $\mathcal{F}$ are distinct, and the objective is the sum of service costs in $C$, i.e., $\sum_{j \in C} d(j, F)$. Chen [13] gives the first constant-factor approximation for R$k$Med. Krishnaswamy et al. [23] employ an iterative rounding method and obtain an approximation ratio of $(7.081+\epsilon)$ for R$k$Med, which is later improved to $(6.994+\epsilon)$ by Gupta et al. [16]. The objective of R$k$Med is equivalent to $\boldsymbol{\sigma}_{|\mathcal{C}|-m+1}^{\top} \boldsymbol{c}_0^{\downarrow}$, where $\boldsymbol{\sigma}_t = \{0, \ldots, 0, 1, \ldots, 1\}$ is the all-one vector except for its first $(t-1)$ coordinates, which are 0's.

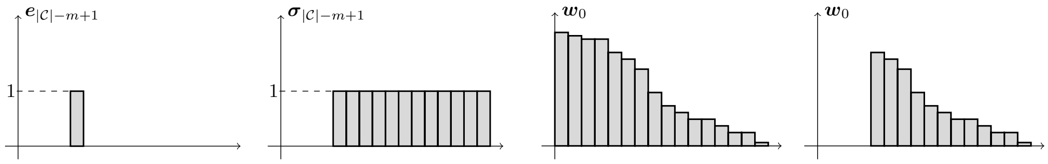
In this paper, we study a new problem called robust ordered $k$-median (RO$k$Med). Formally, given the basic input $(\mathcal{F}, \mathcal{C}, d, k)$, a parameter $m \leq |\mathcal{C}|$ and a non-increasing non-negative vector $\boldsymbol{w} \in \mathbb{R}_+^m$, we are asked to open $k$ facilities $F \subseteq \mathcal{F}$ and serve $m$ clients $C \subseteq \mathcal{C}$, inducing a service cost vector $\boldsymbol{c} = \{d(j, F) : j \in C\} \in \mathbb{R}_+^C$ (notice that $\boldsymbol{c}$ is different from $\boldsymbol{c}_0$, since $\boldsymbol{c}_0$ is indexed by $\mathcal{C}$); the goal is to minimize $\boldsymbol{w}^{\top} \boldsymbol{c}^{\downarrow}$. Clearly, RO$k$Med unifies the aforementioned problems of O$k$Med, R$k$Cen and R$k$Med by choosing $m$ and $\boldsymbol{w}$ suitably.

We can also define the objective of RO$k$Med using $\boldsymbol{c}_0 = \{d(j, F) : j \in \mathcal{C}\}$ as follows. Let $\boldsymbol{w}_0 \in \mathbb{R}_+^{|\mathcal{C}|}$ be a non-negative vector, such that its first $(|\mathcal{C}| - m)$ coordinates are 0's, and the remaining coordinates are non-increasing; the objective of RO$k$Med is $\boldsymbol{w}_0^{\top} \boldsymbol{c}_0^{\downarrow}$. We notice that this weight vector $\boldsymbol{w}_0$ exhibits a distinctly *unimodal* shape; that is, there exists an index $t$ (which is $(|\mathcal{C}| - m + 1)$ here) such that $\boldsymbol{w}_0$ is non-*de*creasing on indexes $\{1, \ldots, t\}$ and non-*in*creasing on indexes $\{t, \ldots, |\mathcal{C}|\}$ (see Figure 1 for an example). Therefore, this objective function places a heavier emphasis on clients that are close to the "mode" of $\boldsymbol{w}_0$. This can also be motivated by the following real-world scenario. Suppose the underlying metric $d$ models the latencies of an online streaming service in different regions (i.e., clients), where the facilities represent potential data center locations. From a business point of view, one could strategically disregard clients that have very poor latencies (they might stop using the service anyway) *and* clients that have very good latencies (they typically have relatively few issues or complaints); the majority of maintenance and servicing costs will then come from clients with medium latencies. By choosing $\boldsymbol{w}_0$ properly, the objective of RO$k$Med can be the sum of *sorted* service costs, say, between the 35th percentile and the 65th percentile, thus acting as a good optimization objective for this scenario. We believe this motivating example for RO$k$Med offers a practical clustering criterion, and our results will stimulate more studies towards clustering objectives with *arbitrary* unimodal weight vectors.

## 1.1 Our Contributions

We first study robust ordered $k$-median and obtain the following main result of this paper.

▶ **Theorem 1.** *There exists a polynomial-time* 127*-approximation algorithm for* RO$k$Med.




**(a)** R$k$Cen. **(b)** R$k$Med. **(c)** O$k$Med. **(d)** RO$k$Med.

**Figure 1** An illustration of different weight vectors used in the objectives of R$k$Cen, R$k$Med, O$k$Med and RO$k$Med. The coordinates are represented using unit-width rectangles.

At a high level, we build a simple reduction framework that reduces each RO$k$Med instance $\mathscr{I}$ to an instance $\mathscr{J}$ of a new problem; the objective of $\mathscr{J}$ is still non-linear, but is formulated as a simple sum and easier to approximate. Moreover, by (approximately) solving the new problem, we show that the approximation guarantee of the solution in $\mathscr{J}$ is preserved up to a constant factor in $\mathscr{I}$ (see Theorem 3 for the formal statement). Thus, it suffices to obtain a constant-factor approximate solution for each new instance $\mathscr{J}$. To this end, we adapt the iterative rounding algorithm by Krishnaswamy et al. [23]. We note that this rounding algorithm is only applicable to the non-linear objective of $\mathscr{J}$ thanks to our *parameterized* reduction framework. Though Gupta et al. [16] give a slightly improved iterative rounding algorithm, we do not adapt their algorithm here. We choose the original algorithm in [23] for its simplicity of presentation. The improvement based on [16] is likely to be small due to our different metric discretization method.

We extend our results to ordered matroid median (OMatMed) and ordered knapsack median (OKnapMed), which are natural generalizations of O$k$Med by replacing the cardinality constraint $|F| \leq k$ with a matroid constraint and a knapsack constraint, respectively (see Section 3.2 for the formal definitions). To the best of our knowledge, no approximation algorithms are known for OMatMed and OKnapMed prior to our study.

▶ **Theorem 2.** *There exist a polynomial-time* 19.8*-approximation algorithm for* OMatMed *and a polynomial-time* 41.6*-approximation algorithm for* OKnapMed*.*

## 1.2 Overview of Techniques

Above all, we need to have apt approximate forms of the ordered objective and write a suitable LP relaxation for RO$k$Med. To start with, let us first review the sparsification method proposed by Aouad and Segev [1] and Byrka et al. [5] for O$k$Med. In the pre-processing phase, one first guesses disjoint intervals $I_0, I_1, \ldots$ with each $I_t \subseteq \mathbb{R}_+$ having the form $(x, (1+\epsilon)x]$ for some small $\epsilon > 0$, so that the service costs falling into the same interval differ by only a multiplicative factor of $(1+\epsilon)$. Let $\boldsymbol{w}_I^{\mathsf{avg}}$ be the average weight multiplied with service costs in the interval $I$ in a fixed optimal solution. If we apply the same weight $\boldsymbol{w}_I^{\mathsf{avg}}$ to all service costs in $I$, we can show that the optimal solution exhibits a similar objective by only losing a $(1+O(\epsilon))$ factor. The pre-processing phase proceeds to build the premise that the guessed intervals $\{I_0, I_1, \ldots\}$ and the guessed average weights $\{\boldsymbol{w}_{I_0}^{\mathsf{avg}}, \boldsymbol{w}_{I_1}^{\mathsf{avg}}, \ldots\}$ roughly agree with the unknown optimal solution; this is done by showing the number of necessary guesses is bounded by a polynomial, thus we can use exhaustive search. To "pre-apply" the average weights in an LP relaxation, we define a function $f$ as $f(d(i,j)) = \boldsymbol{w}_I^{\mathsf{avg}} \cdot d(i,j)$ for $d(i,j) \in I$, and put $f(d(i,j))$ instead of $d(i,j)$ in the LP objective. Byrka et al. [5] implicitly use such a function and give a $(38+\epsilon)$-approximation for O$k$Med.

Unfortunately for RO$k$Med, this objective function seems to suffer from the inherent unbounded integrality gap in the natural relaxation for basic R$k$Med (see, e.g., [23]; also recall that R$k$Med is a special case of RO$k$Med). Note that this is not an issue for O$k$Med since the integrality gap in the natural relaxation for $k$-median is a constant [11, 20]. Roughly speaking, to overcome this gap in the robust case and obtain constant-factor approximation guarantees, one usually strengthens the relaxation by adding more constraints, obtains an almost-integral solution via an auxiliary LP, and rounds the last few fractionally-open facilities to integral ones. During the last step, extra facility-client connections will incur extra costs; to the best of our efforts, the non-linearity of the ordered objective prevents us from obtaining a constant-factor approximate solution, even if we use the aforementioned new LP objective defined via $f$.

We overcome this technical barrier by considering another simple but effective objective function. We replace $f(d(i,j))$ with $f(\lambda d(i,j))$, where $\lambda \in (0,1]$ is a small constant parameter and $f$ is defined similarly as above. We note that the same function has been used in [1] to provide a *logarithmic* approximation guarantee for O$k$Med; we give a much tighter analysis here and achieve a *constant* guarantee. Intuitively speaking, by scaling the underlying metric and still comparing the solutions with the original optimum, we can bound the extra costs incurred in the robust case. We point out that for the optimal service cost vector $\boldsymbol{o}$, the gap between each $f(\boldsymbol{o}_j)$ and $f(\lambda \boldsymbol{o}_j)$ may be $\omega(1/\lambda)$, since $f$ is in fact non-decreasing and superlinear. Nevertheless, we overcome this new gap by obtaining a linear upper bound for any integral solution to the new relaxation. More specifically, let $\mathsf{opt} \geq 0$ be the optimum of the original instance; we show that any integral solution with an objective of $V$ in the $\lambda$-scaled relaxation induces a solution to the original problem with an objective of at most $\lambda^{-1}(V + O(1)\mathsf{opt})$. Furthermore, we show that there exists an algorithm which outputs an integral solution with an objective of $V = O(\lambda)\mathsf{opt}$. Combining these two results, we obtain an approximate solution for RO$k$Med with an objective that is $O(1/\lambda)$ times the optimum.

### 1.3 Other Related Work

Clustering problems with more general combinatorial constraints have been extensively studied in recent years. Chen et al. [12] give a 3-approximation for matroid center. Krishnaswamy et al. [22] give the first constant-factor approximation for matroid median, and thereafter the ratio is improved in [11, 27]; currently the best ratio is 7.081 due to Krishnaswamy et al. [23]. Hochbaum and Shmoys [18] study knapsack center and give a 3-approximation. As for knapsack median, Kumar [24] gives the first constant-factor approximation algorithm; the ratio is later improved in [3, 11, 23, 27], and the best ratio so far is $(6.387 + \epsilon)$ due to Gupta et al. [16].

## 2 The Reduction Framework

In this section, we maintain a generic problem called OrdClst, i.e., ordered clustering, which can be later instantiated as different concrete problems such as RO$k$Med. An instance $\mathscr{I}$ of OrdClst consists of a facility set $\mathcal{F}$, a client set $\mathcal{C}$, a finite metric $d$ on $\mathcal{F} \cup \mathcal{C}$, feasible facility sets $\mathscr{F} \subseteq 2^{\mathcal{F}}$, feasible client sets $\mathscr{C} \subseteq 2^{\mathcal{C}}$, and a non-increasing non-negative vector $\boldsymbol{w} \in \mathbb{R}_+^m$; each $C \in \mathscr{C}$ satisfies $|C| = m$, and $d(u,v) \geq 1$ for $u, v \in \mathcal{F} \cup \mathcal{C}$ that are *not* co-located. The goal is to choose $F \in \mathscr{F}$ and $C \in \mathscr{C}$ that induce a service cost vector $\boldsymbol{c} = \{d(j,F) : j \in C\}$ such that the ordered objective $cost(\boldsymbol{w}; \boldsymbol{c}) = \boldsymbol{w}^\top \boldsymbol{c}^\downarrow$ is minimized.

We devise a general framework that reduces OrdClst instances to other clustering problems with simpler objective functions. Given an instance $\mathscr{I} = (\mathcal{F}, \mathcal{C}, d, \mathscr{F}, \mathscr{C}, \boldsymbol{w})$ of OrdClst and a *non-decreasing* function $f : \mathbb{R}_+ \to \mathbb{R}_+$, we say $\mathscr{J} = (\mathcal{F}, \mathcal{C}, d, \mathscr{F}, \mathscr{C}, f)$ is a *reduced instance* of $\mathscr{I}$, whose goal of optimization is to choose $F \in \mathscr{F}$ and $C \in \mathscr{C}$ such that the new objective $\sum_{j \in C} f(d(j,F))$ is minimized. Using this reduction, we will show that when $f$ satisfies some certain nice properties, we only need to study the reduced instance $\mathscr{J}$, whose objective might be more tangible and easier to deal with. Moreover, we will show that an approximate solution to the original instance $\mathscr{I}$ can be directly recovered from an approximate solution to $\mathscr{J}$ by only losing a constant factor in the approximation guarantee.

The framework adapts previous sparsification methods [1, 5] for O$k$Med, and generalizes the helper functions therein to overcome the technical difficulties that may be present in OrdClst (see Section 1.2 for the discussion). For convenience, for each function $f:\mathbb{R}_+\to\mathbb{R}_+$ and $\lambda>0$, we define $f_\lambda(x):=f(\lambda x)$, $\forall x\geq 0$. We let $n_0=|\mathcal{F}\cup\mathcal{C}|$ and present the following core theorem of the reduction framework.

▶ **Theorem 3.** *Let $\mathscr{I}=(\mathcal{F},\mathcal{C},d,\mathscr{F},\mathscr{C},\boldsymbol{w})$ be an instance of OrdClst with optimum* $\mathsf{opt}\geq 0$. *For each $\epsilon\in(0,1)$, there exists an algorithm that outputs $(n_0/\epsilon)^{O(1/\epsilon)}$ non-decreasing functions $\mathbb{R}_+\to\mathbb{R}_+$, such that there is an output $f$ satisfying the following for each $\lambda\in(0,1]$. We say such $f$ is* faithful.

- *The reduced instance $\mathscr{J}_{f_\lambda}=(\mathcal{F},\mathcal{C},d,\mathscr{F},\mathscr{C},f_\lambda)$ has an optimum of at most $\lambda(1+9\epsilon)\mathsf{opt}$.*
- *If an algorithm produces a solution with objective $V$ for $\mathscr{J}_{f_\lambda}$, the same solution attains an objective of at most $\lambda^{-1}(V+(1+4\epsilon)\mathsf{opt})$ for $\mathscr{I}$.*

As a direct consequence, if we obtain a solution to any "faithfully" reduced instance $\mathscr{J}_{f_\lambda}$ with an objective of $V\leq\gamma\mathsf{opt}$, it is a $\lambda^{-1}(1+\gamma+4\epsilon)$-approximate solution to $\mathscr{I}$. Before we proceed with the proof, we discuss the sparsification method used for constructing such functions. We shall only consider these functions in the remainder of this section. We note that the same functions are also used in a much more straightforward fashion in [5].

Let $(F^\star,C^\star)$ be a fixed (unknown) optimal solution to the original problem, $\boldsymbol{o}\in\mathbb{R}_+^{C^\star}$ be the corresponding service cost vector, and $\mathsf{opt}=\mathit{cost}(\boldsymbol{w};\boldsymbol{o})$ be the optimal objective thereof. We first guess the exact value of $\boldsymbol{o}_1^\downarrow$, i.e., the largest service cost, which only has a polynomial number of possible values. We use exhaustive search and assume $\boldsymbol{o}_1^\downarrow$ is correctly guessed in the sequel; we also assume $\boldsymbol{o}_1^\downarrow>0$, otherwise the solution is trivial.

Let $T$ be the smallest integer s.t. $\epsilon(1+\epsilon)^T>m$ and define intervals $I_{T+1},I_T,\ldots,I_0$ where

$$I_{T+1}=\left[0,\frac{\epsilon\boldsymbol{o}_1^\downarrow}{m}\right];\ I_t=\left(\frac{\epsilon\boldsymbol{o}_1^\downarrow}{m}(1+\epsilon)^{T-t},\frac{\epsilon\boldsymbol{o}_1^\downarrow}{m}(1+\epsilon)^{T-t+1}\right],\ \forall t\in[T];\ I_0=\left(\frac{\epsilon\boldsymbol{o}_1^\downarrow}{m}(1+\epsilon)^T,+\infty\right).$$

Since $\bigcup_{t=0}^{T+1}I_t=\mathbb{R}_+$ and they are mutually disjoint, each $d(i,j)$ falls into exactly one interval.

Next, to avoid technical difficulties caused by weights that are too small, we define a new vector $\widetilde{\boldsymbol{w}}$ where $\widetilde{\boldsymbol{w}}_i=\max\{\boldsymbol{w}_i,\frac{\epsilon\boldsymbol{w}_1}{m}\}$, $i\in[m]$. We obtain the following simple fact, by observing $\widetilde{\boldsymbol{w}}\geq\boldsymbol{w}$ and $\mathit{cost}(\widetilde{\boldsymbol{w}};\boldsymbol{v})-\mathit{cost}(\boldsymbol{w};\boldsymbol{v})\leq m\cdot\frac{\epsilon\boldsymbol{w}_1}{m}\cdot\boldsymbol{v}_1^\downarrow\leq\epsilon\cdot\mathit{cost}(\boldsymbol{w};\boldsymbol{v})$.

▶ **Fact 4.** *For each $\boldsymbol{v}\subseteq\mathbb{R}_+^m$, one has $\mathit{cost}(\boldsymbol{w};\boldsymbol{v})\leq\mathit{cost}(\widetilde{\boldsymbol{w}};\boldsymbol{v})\leq(1+\epsilon)\mathit{cost}(\boldsymbol{w};\boldsymbol{v})$.*

Now, let us consider the optimum $\mathsf{opt}=\boldsymbol{w}^\top\boldsymbol{o}^\downarrow$. In particular, we consider the entries of $\boldsymbol{o}^\downarrow$ that fall into different intervals $I_{T+1},I_T,\ldots,I_0$, and (iteratively) define the average weight $\boldsymbol{w}_t^{\mathsf{avg}}$ w.r.t. $\boldsymbol{o}^\downarrow$, $\widetilde{\boldsymbol{w}}$ and interval $I_t$, such that $\boldsymbol{w}_0^{\mathsf{avg}}=\widetilde{\boldsymbol{w}}_1$ and

$$\boldsymbol{w}_t^{\mathsf{avg}}=\begin{cases}\left(\sum_{j:\boldsymbol{o}_j^\downarrow\in I_t}\widetilde{\boldsymbol{w}}_j\right)\Big/\left|\boldsymbol{o}^\downarrow\cap I_t\right| & \boldsymbol{o}^\downarrow\cap I_t\neq\emptyset,\ t\geq 1,\\ \boldsymbol{w}_{t-1}^{\mathsf{avg}} & \boldsymbol{o}^\downarrow\cap I_t=\emptyset,\ t\geq 1.\end{cases}$$

Since $\widetilde{\boldsymbol{w}}$ is non-increasing, it follows that $\boldsymbol{w}^{\mathsf{avg}}$ is also non-increasing. Though the actual sequence $\boldsymbol{w}^{\mathsf{avg}}$ is unknown, we can estimate it using another non-increasing sequence $\boldsymbol{w}^{\mathsf{gss}}$ such that for each $0\leq t\leq T+1$, $\boldsymbol{w}_t^{\mathsf{gss}}$ is an integer power of $(1+\epsilon)$ and satisfies $\min_s\boldsymbol{w}_s^{\mathsf{avg}}\leq\boldsymbol{w}_t^{\mathsf{gss}}\leq(1+\epsilon)\max_s\boldsymbol{w}_s^{\mathsf{avg}}$. Since the entries of $\boldsymbol{w}^{\mathsf{avg}}$ are at least $\min_j\widetilde{\boldsymbol{w}}_j\geq\epsilon\boldsymbol{w}_1/m$ and at most $\boldsymbol{w}_1$, the number of possible values is $O(\log_{1+\epsilon}(m/\epsilon))$. By the definition of $T$, we have $T = O(\log_{1+\epsilon}(m/\epsilon))$. Thus, using routine calculation, the number of all possible non-increasing sequences for $\boldsymbol{w}^{\mathsf{gss}}$ is at most $(m/\epsilon)^{O(1/\epsilon)}$. Up to now, we have only guessed $\boldsymbol{o}_1^{\downarrow}$ and $\boldsymbol{w}^{\mathsf{gss}}$, hence the total number of possible guesses is at most $(n_0/\epsilon)^{O(1/\epsilon)}$ since $m \le |\mathcal{C}| \le n_0$.

**Proof of Theorem 3.** For each guess $(\boldsymbol{o}_1^{\downarrow}, \{\boldsymbol{w}_t^{\mathsf{gss}}\}_{t=0}^{T+1})$, we define a piece-wise linear function

$$f(x) = \boldsymbol{w}_t^{\mathsf{gss}} x,\ x \in I_t,\ 0 \le t \le T+1.$$

Because $\boldsymbol{w}^{\mathsf{gss}}$ is non-increasing, $f : \mathbb{R}_+ \to \mathbb{R}_+$ is non-decreasing and *superlinear* (i.e., $f(\alpha x) \ge \alpha f(x)$ for each $\alpha \ge 1$ and $x \ge 0$). According to the previous analysis, we consider at most $(n_0/\epsilon)^{O(1/\epsilon)}$ such functions.

To prove the theorem, it suffices to show the existence of a faithful function. In the sequel, we assume that the guessed values are as desired; that is, $\boldsymbol{o}_1^{\downarrow}$ is precisely the largest service cost in the optimal solution and for each $0 \le t \le T+1$, one has $\boldsymbol{w}_t^{\mathsf{gss}} \in [\boldsymbol{w}_t^{\mathsf{avg}}, (1+\epsilon)\boldsymbol{w}_t^{\mathsf{avg}})$. We show that the corresponding function $f$ is faithful. We need the following two lemmas.

▶ **Lemma 5.** *Let $\boldsymbol{c} = \{d(j,F) : j \in C\} \in \mathbb{R}_+^C$ where $F \in \mathscr{F}$ and $C \in \mathscr{C}$. For each $\lambda \in (0,1]$, one has $\lambda \cdot \widetilde{\boldsymbol{w}}^\top \boldsymbol{c}^{\downarrow} \le \sum_{j\in C} f(\lambda \boldsymbol{c}_j) + (1+3\epsilon+\epsilon^2)\mathsf{opt}$.*

**Proof.** Recall that $|C| = m$ for each feasible $C \in \mathscr{C}$. We suppose $C = [m]$ for convenience. Consider each $j \in [m]$ s.t. $\lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow} > f(\lambda \boldsymbol{c}_j^{\downarrow})$. Notice that $\lambda \boldsymbol{c}_j^{\downarrow} \notin I_0$, otherwise one has $\boldsymbol{w}_0^{\mathsf{gss}} \ge \boldsymbol{w}_0^{\mathsf{avg}} = \widetilde{\boldsymbol{w}}_1 \ge \widetilde{\boldsymbol{w}}_j$ and $\lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow} > f(\lambda \boldsymbol{c}_j^{\downarrow}) = \boldsymbol{w}_0^{\mathsf{gss}}(\lambda \boldsymbol{c}_j^{\downarrow}) \ge \lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow}$, which is a contradiction. If $\lambda \boldsymbol{c}_j^{\downarrow} \in I_{T+1} = [0, \epsilon \boldsymbol{o}_1^{\downarrow}/m]$, one has $\lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow} \le \widetilde{\boldsymbol{w}}_j(\epsilon \boldsymbol{o}_1^{\downarrow}/m) \le \epsilon \cdot \boldsymbol{w}^\top \boldsymbol{o}^{\downarrow}/m$ since $\widetilde{\boldsymbol{w}}_j \le \widetilde{\boldsymbol{w}}_1 = \boldsymbol{w}_1$.

Then, suppose $\lambda \boldsymbol{c}_j^{\downarrow} \in I_t$, $t \in [T]$. We claim $\lambda \boldsymbol{c}_j^{\downarrow} \le (1+\epsilon)\boldsymbol{o}_j^{\downarrow}$. For the sake of contradiction, assume otherwise, i.e., $\lambda \boldsymbol{c}_j^{\downarrow} > (1+\epsilon)\boldsymbol{o}_j^{\downarrow}$, thus $\lambda \boldsymbol{c}_j^{\downarrow}$ and $\boldsymbol{o}_j^{\downarrow}$ must be in different intervals by the definition of $I_t$. Suppose $\boldsymbol{o}_j^{\downarrow} \in I_{t'}$ for some $t' > t$, which implies $\boldsymbol{w}_t^{\mathsf{avg}} \ge \widetilde{\boldsymbol{w}}_j$, because $\boldsymbol{w}_t^{\mathsf{avg}}$ is the average weight on $I_t$ w.r.t. $\boldsymbol{o}^{\downarrow}$, and $\widetilde{\boldsymbol{w}}_j$ is the weight for $\boldsymbol{o}_j^{\downarrow} \in I_{t'}$. Therefore, because $\boldsymbol{w}_t^{\mathsf{gss}} \ge \boldsymbol{w}_t^{\mathsf{avg}}$ using our initial conditions, we have $f(\lambda \boldsymbol{c}_j^{\downarrow}) = \boldsymbol{w}_t^{\mathsf{gss}}(\lambda \boldsymbol{c}_j^{\downarrow}) \ge \lambda \boldsymbol{w}_t^{\mathsf{avg}} \boldsymbol{c}_j^{\downarrow} \ge \lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow}$, contradicting our initial assumption. Thus the claim is true.

The above analysis shows that $\lambda \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow} \le f(\lambda \boldsymbol{c}_j^{\downarrow}) + (1+\epsilon)\widetilde{\boldsymbol{w}}_j \boldsymbol{o}_j^{\downarrow} + \epsilon \boldsymbol{w}^\top \boldsymbol{o}^{\downarrow}/m$ for each $j \in [m]$. We take the sum over $j \in [m]$ and obtain

$$\lambda \cdot \widetilde{\boldsymbol{w}}^\top \boldsymbol{c}^{\downarrow} = \lambda \sum_{j\in[m]} \widetilde{\boldsymbol{w}}_j \boldsymbol{c}_j^{\downarrow} \le \sum_{j\in[m]} f(\lambda \boldsymbol{c}_j^{\downarrow}) + (1+\epsilon) cost(\widetilde{\boldsymbol{w}}; \boldsymbol{o}) + \epsilon \cdot cost(\boldsymbol{w}; \boldsymbol{o}).$$

Combining with Fact 4 and $\mathsf{opt} = cost(\boldsymbol{w}; \boldsymbol{o})$, the lemma follows. ◀

▶ **Lemma 6.** *For each $\lambda \in (0,1]$, one has $\sum_{j\in C^\star} f(\lambda \boldsymbol{o}_j) \le \lambda((1+\epsilon)^3 + \epsilon + \epsilon^2)\mathsf{opt}$.*

**Proof.** Recall that $(F^\star, C^\star)$ is optimal for $\mathscr{I}$. Consider any non-empty $\boldsymbol{o} \cap I_t$, and it is easy to verify that $t > 0$. Since $\lambda \le 1$, some entries in $\lambda(\boldsymbol{o} \cap I_t)$ may be "shifted" to $I_{t'}$ with $t' > t$. If $t \le T$, the contribution of $\lambda(\boldsymbol{o} \cap I_t)$ on the LHS is at most

$$\sum_{\substack{j:\boldsymbol{o}_j^{\downarrow}\in I_t,\\ \lambda \boldsymbol{o}_j^{\downarrow}\notin I_t}} \lambda \boldsymbol{w}_{t+1}^{\mathsf{gss}} \boldsymbol{o}_j^{\downarrow} + \sum_{\substack{j:\boldsymbol{o}_j^{\downarrow}\in I_t,\\ \lambda \boldsymbol{o}_j^{\downarrow}\in I_t}} \lambda \boldsymbol{w}_t^{\mathsf{gss}} \boldsymbol{o}_j^{\downarrow} \le \lambda \sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t} (1+\epsilon)\boldsymbol{w}_t^{\mathsf{avg}} \boldsymbol{o}_j^{\downarrow} \le \lambda(1+\epsilon)^2 \sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t} \widetilde{\boldsymbol{w}}_j \boldsymbol{o}_j^{\downarrow}, \quad (1)$$

where the first inequality is due to non-increasing $\boldsymbol{w}^{\mathsf{gss}}$ and $\boldsymbol{w}_t^{\mathsf{gss}} \in [\boldsymbol{w}_t^{\mathsf{avg}}, (1+\epsilon)\boldsymbol{w}_t^{\mathsf{avg}})$; the second inequality is because within the same interval $I_t$ where $t \le T$, the values of $\boldsymbol{o}_j^{\downarrow}$ differ by a factor no more than $(1+\epsilon)$. More formally, we have

$$\sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t} \boldsymbol{w}_t^{\mathsf{avg}}\boldsymbol{o}_j^{\downarrow} = \left(\frac{1}{|\boldsymbol{o}\cap I_t|}\sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t}\widetilde{\boldsymbol{w}}_j\right)\sum_{j':\boldsymbol{o}_{j'}^{\downarrow}\in I_t}\boldsymbol{o}_{j'}^{\downarrow}$$
$$= \sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t}\left(\frac{1}{|\boldsymbol{o}\cap I_t|}\sum_{j':\boldsymbol{o}_{j'}^{\downarrow}\in I_t}\boldsymbol{o}_{j'}^{\downarrow}\right)\widetilde{\boldsymbol{w}}_j \le \sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t}(1+\epsilon)\boldsymbol{o}_j^{\downarrow}\widetilde{\boldsymbol{w}}_j,$$

hence the inequality above follows.

If $t = T+1$, each such $\boldsymbol{o}_j^{\downarrow} \le \epsilon\boldsymbol{o}_1^{\downarrow}/m$, thus the contribution of $\lambda(\boldsymbol{o}\cap I_{T+1})$ is at most $\epsilon\lambda\boldsymbol{w}_1^{\mathsf{gss}}\boldsymbol{o}_1^{\downarrow} \le \epsilon(1+\epsilon)\lambda\widetilde{\boldsymbol{w}}_1\boldsymbol{o}_1^{\downarrow} \le \lambda(\epsilon+\epsilon^2)\mathsf{opt}$, since $\boldsymbol{w}_1^{\mathsf{gss}} \le (1+\epsilon)\boldsymbol{w}_1^{\mathsf{avg}} \le (1+\epsilon)\widetilde{\boldsymbol{w}}_1$. The lemma follows by taking the sum of (1) over each $\boldsymbol{o}\cap I_t$ plus $\lambda(\epsilon+\epsilon^2)\mathsf{opt}$ for $\boldsymbol{o}\cap I_{T+1}$, which is

$$\sum_{j\in C^\star} f(\lambda\boldsymbol{o}_j) \le \lambda(1+\epsilon)^2\sum_{t=1}^{T}\sum_{j:\boldsymbol{o}_j^{\downarrow}\in I_t}\widetilde{\boldsymbol{w}}_j\boldsymbol{o}_j^{\downarrow} + \lambda(\epsilon+\epsilon^2)\mathsf{opt}$$
$$\stackrel{\text{(Fact 4)}}{\le} \lambda(1+\epsilon)^3 cost(\boldsymbol{w};\boldsymbol{o}) + \lambda(\epsilon+\epsilon^2)\mathsf{opt}. \qquad \blacktriangleleft$$

We return to the original theorem and fix $\lambda \in (0,1]$. Since $(F^\star, C^\star)$ is a feasible solution to both $\mathscr{I}$ and $\mathscr{I}_{f_\lambda}$, the first assertion follows using $\epsilon < 1$ and Lemma 6. For the second assertion, let $(F, C)$ be the solution returned by the algorithm, thus $V = \sum_{j\in C} f(\lambda d(j,F))$. Therefore, using Fact 4 and Lemma 5, the objective of $(F,C)$ in the OrdClst instance $\mathscr{I}$ is at most $cost(\boldsymbol{w};\boldsymbol{c}) \le cost(\widetilde{\boldsymbol{w}};\boldsymbol{c}) \le \lambda^{-1}(V + (1+4\epsilon)\mathsf{opt})$, where $\boldsymbol{c} = \{d(j,F) : j \in C\}$. ◀

# 3 Applications

In this section, we provide applications of our reduction framework in Theorem 3. Due to the space limitations, we defer some proofs and details of the algorithms to the appendix.

## 3.1 Robust Ordered $k$-Median

In RO$k$Med, OrdClst is instantiated such that $\mathscr{F}$ consists of all subsets of $\mathcal{F}$ with cardinality at most $k$, i.e., $\mathscr{F} = \{F \subseteq \mathcal{F} : |F| \le k\}$; $\mathscr{C}$ consists of all subsets of $\mathcal{C}$ with cardinality exactly $m$, i.e., $\mathscr{C} = \{C \subseteq \mathcal{C} : |C| = m\}$. Via enumerating all possible functions in Theorem 3, suppose that we have a faithful function $f$ in what follows.

As noted before, using Theorem 3, we want to obtain a constant-factor approximate solution to $\mathscr{I}_{f_\lambda}$ for some small constant $\lambda \in (0,1)$. We adapt the iterative rounding algorithm [23]. Let $x_{ij} \in [0,1]$ denote the extent of assigning client $j$ to facility $i$, and $y_i \in [0,1]$ denote the extent of opening facility $i$. The natural relaxation for $\mathscr{I}_{f_\lambda}$ is given as follows.

$$\begin{aligned}
\min \quad & \sum_{j\in\mathcal{C}}\sum_{i\in\mathcal{F}} x_{ij} f_\lambda(d(i,j)) && \text{(LP}(f_\lambda))\\
\text{s.t.} \quad & \sum_{j\in\mathcal{C}}\sum_{i\in\mathcal{F}} x_{ij} \ge m \\
& \sum_{i\in\mathcal{F}} x_{ij} \le 1 && \forall j \in \mathcal{C}\\
& \sum_{i\in\mathcal{F}} y_i \le k \\
& 0 \le x_{ij} \le y_i \le 1 && \forall i \in \mathcal{F}, j \in \mathcal{C}.
\end{aligned}$$

Before we look at the full algorithm, let us begin with a brief overview. The algorithm consists of the following two phases, namely, pre-processing and iterative rounding.

**Pre-processing.** As is discussed in the introduction, the integrality gap in LP($f_\lambda$) is unbounded since RO$k$Med recovers R$k$MED. Thus, instead of directly solving LP($f_\lambda$), we employ some pre-processing techniques and simplify the instance. In what follows, let $\lambda_1 \in (\lambda, 1]$ be another constant. The values of $\lambda$ and $\lambda_1$ will be determined in the full algorithm.

First, we guess a constant number of facilities $S_0$ as must-have choices and remove some clients in advance. Consequently, we obtain a new *extended* instance $\mathscr{I}'_f$ on the remaining clients $\mathcal{C}'$; the family $\mathscr{C}' \subseteq 2^{\mathcal{C}'}$ of client subsets in $\mathscr{I}'_f$ consists of all $m'$-subsets of $\mathcal{C}'$, that is, $\mathscr{C}' = \{C' \subseteq \mathcal{C}' : |C'| = m'\}$ for some fixed parameter $m' \leq m$; $\mathscr{I}'_f$ also requires that the pre-selected facilities in $S_0$ must be part of the solution. By exhaustive search, we show that there exists some $\mathscr{I}'_f$ with certain "sparse" properties, which is easier to approximate using an LP relaxation. More specifically, there exists a solution $(F, C)$ to $\mathscr{I}'_f$ such that $S_0 \subseteq F$ and $(F, C)$ satisfies the following for two small constants $\rho, \delta \in (0, 1)$ (see Theorem 7). Here, $\mathsf{opt} \geq 0$ is the optimum of the original RO$k$Med instance.

- For each facility $i \in F \setminus S_0$, the clients assigned to $i$ contribute at most $\rho \cdot \mathsf{opt}$; that is, $\sum_{j \in C \text{ assigned to } i} f(d(i,j)) \leq \rho \cdot \mathsf{opt}$.
- For each $p \in \mathcal{F} \cup \mathcal{C}'$, let $c_p = d(p, F)$. The product of (a) $f((1-\delta)c_p)$ and (b) the number of served clients in $C$ within a distance of $\delta c_p$ from $p$ is at most $\rho \cdot \mathsf{opt}$; that is,

$$|\{j \in C : d(j,p) \leq \delta c_p\}| \cdot f((1-\delta)c_p) \leq \rho \cdot \mathsf{opt}.$$

  Intuitively speaking, this means that after removing the clients $\mathcal{C} \setminus \mathcal{C}'$, there cannot be too many clients with "large" contributions inside any such closed ball.

Further, a straightforward greedy algorithm on the removed clients $\mathcal{C} \setminus \mathcal{C}'$ recovers a good approximate solution to $\mathscr{I}_f$. The two basic instances $\{\mathscr{I}'_f, \mathscr{I}_f\}$ will be useful in the analysis.

Second, we formulate a stronger relaxation S-LP($f_{\lambda_1}$) that has the same objective as LP($f_\lambda$) except for using a larger coefficient $\lambda_1$. It has both the constraints of LP($f_\lambda$), and additional constraints that guarantee certain sparse properties in its solutions; in particular, $\mathscr{I}'_f$ also conforms to these sparsity constraints. Thus, we show that any sparse solution to $\mathscr{I}'_{f_{\lambda_1}}$ is also feasible to S-LP($f_{\lambda_1}$). We emphasize that during the algorithm, we solve S-LP($f_{\lambda_1}$) instead of LP($f_\lambda$); LP($f_\lambda$) will only be used in the analysis of the algorithm.

**Iterative rounding.** After obtaining a fractional optimal solution to S-LP($f_{\lambda_1}$), we use the iterative rounding algorithm and obtain an integral solution $(\hat{F}, \hat{C}')$ to $\mathscr{I}'_{f_{\lambda_1}}$. As aforementioned, it is easy to extend $(\hat{F}, \hat{C}')$ to another solution $(\hat{F}, \hat{C})$ that is feasible to $\mathscr{I}_{f_{\lambda_1}}$. However, because the function $f$ in the LP objective is superlinear and our rounding algorithm incurs multiplicative factors on the input of $f$, we cannot directly analyze the approximation guarantee via $\mathscr{I}_{f_{\lambda_1}}$ and S-LP($f_{\lambda_1}$). Nevertheless, $(\hat{F}, \hat{C})$ is also feasible to LP($f_\lambda$) where the coefficient $\lambda$ is smaller than $\lambda_1$, making it possible for us to bound the objective of $(\hat{F}, \hat{C})$ in the instance $\mathscr{I}_{f_\lambda}$. Finally, we invoke Theorem 3 on $\mathscr{I}_{f_\lambda}$ and obtain the overall approximation ratio.

### 3.1.1 The Algorithm for Robust Ordered $k$-Median

In this section, we present our constant-factor approximation algorithm for RO$k$Med and prove Theorem 1. Suppose we have a faithful function $f : \mathbb{R}_+ \to \mathbb{R}_+$ via Theorem 3 and exhaustive search. Let the reduced instance be $\mathscr{I} = (\mathcal{F}, \mathcal{C}, d, \mathscr{F}, \mathscr{C}, f)$. Recall $n_0 = |\mathcal{F} \cup \mathcal{C}|$.

#### 3.1.1.1 The Sparse Instance

Let $(F^\star \in \mathscr{F}, C^\star \in \mathscr{C})$ be a fixed *unknown* optimal solution to the original RO$k$Med instance $\mathscr{I} = (\mathcal{F}, \mathcal{C}, d, \mathscr{F}, \mathscr{C}, \boldsymbol{w})$ and $\mathsf{opt} \geq 0$ be the optimum thereof; define $c_p^\star = \min_{i \in F^\star} d(p,i)$, $\kappa_p^\star = \arg\min_{i \in F^\star} d(p,i)$ for each $p \in \mathcal{F} \cup \mathcal{C}$ (ties broken arbitrarily), and closed balls $\mathsf{Ball}_S(p,R) = \{i \in S : d(i,p) \leq R\}$. We guess $U \in [V^\star, (1+\epsilon)V^\star)$ via binary search, where $V^\star$ is the optimum of $\mathscr{J}$. We have $V^\star \leq (1+O(\epsilon))\mathsf{opt}$ using Theorem 3. We need the following theorems on pre-processing. The proofs are given in the appendix.

▶ **Theorem 7.** (Similar to [23]). *Given $\rho, \delta \in (0,1)$ and $U$, there exists an $n_0^{O(1/\rho)}$-time algorithm that finds an* extended *instance $\mathscr{J}' = (\mathcal{F}, \mathcal{C}', d, \mathscr{F}, \mathscr{C}', f, S_0)$ satisfying the following.*

**(7.1)** *$\mathcal{C}' \subseteq \mathcal{C}$, $\mathscr{C}' = \{C' \subseteq \mathcal{C}' : |C'| = m' := |C^\star \cap \mathcal{C}'|\}$ and $S_0 \subseteq F^\star$ with $|S_0| = O(1/\rho)$.*

**(7.2)** *Denote $C'^\star = C^\star \cap \mathcal{C}'$. For each $i \in F^\star \setminus S_0$, we have $\sum_{j \in C'^\star : \kappa_j^\star = i} f(c_j^\star) \leq \rho U$.*

**(7.3)** *For each $p \in \mathcal{F} \cup \mathcal{C}'$, we have $\left|\mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)\right| \cdot f((1-\delta)c_p^\star) \leq \rho U$.*

**(7.4)** *Denote $U' = \sum_{j \in C'^\star} f(c_j^\star)$. We have $\sum_{j \in C^\star \setminus \mathcal{C}'} f\left(\frac{1-\delta}{1+\delta} d(j, S_0)\right) + U' \leq U$.*

Roughly speaking, Theorem 7 says that after removing a constant number of facilities $S_0$ from $F^\star$ and some clients $\mathcal{C} \setminus \mathcal{C}'$ from $C^\star$, the remaining solution has some nice sparse properties. Moreover, we can easily extend a solution on $\mathscr{J}'$ to another solution on $\mathscr{J}$ using (7.4) such that the objective can still be bounded in terms of $U \leq (1+O(\epsilon))\mathsf{opt}$.

▶ **Theorem 8.** (Similar to [23]). *Given the instance $\mathscr{J}'$ found in Theorem 7, we can efficiently compute a set of upper bounds $\{\widehat{R}_j \geq 0 : j \in \mathcal{C}'\}$ satisfying the following.*

**(8.1)** *There exists a solution $(F^\star, C')$ to $\mathscr{J}'$, such that each $j \in C'$ is assigned to $\kappa_j' \in F^\star$ and $c_j' := d(\kappa_j', j) \leq (1 + 3\delta/4)\widehat{R}_j$. Moreover, one has*

$$\sum_{j \in C'} f\left(\frac{2}{2+\delta} c_j'\right) \leq U'; \quad \sum_{j \in C' : \kappa_j' = i} f\left(\frac{2}{2+\delta} c_j'\right) \leq \rho U, \forall i \in F^\star \setminus S_0.$$

**(8.2)** *For each $t > 0$ and $p \in \mathcal{F} \cup \mathcal{C}'$, one has*

$$\left|\left\{j \in \mathsf{Ball}_{C'}\left(p, \frac{\delta}{4}t\right) : \widehat{R}_j \geq t\right\}\right| \leq \frac{\rho U}{f((1-\delta)(1-\delta/4)t)}.$$

Roughly speaking, Theorem 8 says that we can efficiently find an upper bound $\widehat{R}_j$ for each $j \in \mathcal{C}'$ such that there exists a solution for $\mathscr{J}'$ that roughly respects these upper bounds and exhibits a similar sparse property as (7.2). Moreover, (8.2) is a stronger but somewhat different version of (7.3); its parameterized form will be useful in the analysis of the approximation guarantee.

#### 3.1.1.2 The Strengthened LP

Let $R_j = (1 + 3\delta/4)\widehat{R}_j$ in Theorem 8 and define the following stronger LP relaxation for $0 < \lambda_1 \leq 2/(2+\delta)$. We note that S-LP$(f_{\lambda_1})$ is built on the new instance $\mathscr{J}'$, hence admits a more "regular" solution according to Theorem 8. In our algorithm, we solve S-LP$(f_{\lambda_1})$

instead of LP($f_\lambda$), and conduct iterative rounding on its solution.

$$\begin{aligned}
\min \quad & \sum_{j\in\mathcal{C}'}\sum_{i\in\mathcal{F}} x_{ij} f_{\lambda_1}(d(i,j)) && && (\text{S-LP}(f_{\lambda_1}))\\
\text{s.t.} \quad & \sum_{j\in\mathcal{C}'}\sum_{i\in\mathcal{F}} x_{ij} \geq m' && && \\
& \sum_{i\in\mathcal{F}} x_{ij} \leq 1 && \forall j\in\mathcal{C}' && \\
& \sum_{i\in\mathcal{F}} y_i \leq k && && \\
& 0 \leq x_{ij} \leq y_i \leq 1 && \forall i\in\mathcal{F}, j\in\mathcal{C}' && \\
& y_i = 1 && \forall i \in S_0 && (\text{S-LP.5})\\
& x_{ij} = 0 && d(i,j) > R_j && (\text{S-LP.6})\\
& x_{ij} = 0 && \forall i \notin S_0,\ f\left(2d(i,j)/(2+\delta)\right) > \rho U && (\text{S-LP.7})\\
& \sum_{j\in\mathcal{C}'} f\left(2d(i,j)/(2+\delta)\right) x_{ij} \leq \rho U y_i && \forall i\notin S_0. && (\text{S-LP.8})
\end{aligned}$$

▶ **Lemma 9.** *The optimal objective value of* S-LP($f_{\lambda_1}$) *is at most* $\frac{\lambda_1(2+\delta)}{2}U'$.

**Proof.** Using (8.1), there exists an integral solution with an objective of at most $U'$ when $\lambda_1 = 2/(2+\delta)$. For $\lambda_1 \leq 2/(2+\delta)$, the same solution is still feasible because the constraints are independent of $\lambda_1$. For $\alpha \leq 1, z > 0$, we have $f(\alpha z) \leq \alpha f(z)$ because $f$ is non-decreasing and superlinear, thus $\sum_{j\in C'} f(\lambda_1 c'_j) \leq \frac{\lambda_1(2+\delta)}{2}\sum_{j\in C'} f\left(\frac{2}{2+\delta}c'_j\right) \leq \frac{\lambda_1(2+\delta)}{2}U'$. ◀

After we solve S-LP($f_{\lambda_1}$) and obtain an optimal solution $(x^\star, y^\star)$, to eliminate the $x^\star$ variables and work with an auxiliary LP that is purely on the $y^\star$ variables, we need the following lemma due to [23]. Note that this is different from simple facility duplication [11], since we need a certain sparse property of the modified solution (see 5), which helps us bound the additional rounding cost in the final analysis.

▶ **Lemma 10.** *We can add co-located copies to $\mathcal{F}$, create a vector $y^\star \in [0,1]^{\mathcal{F}}$ and define subsets $F_j \subseteq \mathsf{Ball}_{\mathcal{F}}(j, R_j)$ for each client $j \in \mathcal{C}'$, such that the following holds.*

**(10.1)** *$y^\star(F_j) \leq 1$ for each $j\in\mathcal{C}'$ and $\sum_{j\in\mathcal{C}'}\left(\sum_{i\in F_j} y_i^\star\right) \geq m'$.*

**(10.2)** *$\sum_{i\in\mathcal{F}} y_i^\star \leq k$.*

**(10.3)** *For each $i\in S_0$, $\sum_{i' \text{ co-located with } i} y_i^\star = 1$.*

**(10.4)** *$\sum_{j\in\mathcal{C}'}\sum_{i\in F_j} y_i^\star f_{\lambda_1}(d(i,j)) \leq \frac{\lambda_1(2+\delta)}{2}U'$.*

**(10.5)** *For each $i$ not co-located with $S_0$, $\sum_{j\in\mathcal{C}':i\in F_j} f\left(\frac{2}{2+\delta}d(i,j)\right) \leq 2\rho U$.*

**Proof.** We start with an optimal solution $(x^\star, y^\star)$ to S-LP($f_{\lambda_1}$) with objective at most $\frac{\lambda_1(2+\delta)}{2}U'$ according to Lemma 9. To avoid confusion in notation, we create a copy $\mathcal{F}' = \mathcal{F}$, define $F_j = \{i \in \mathcal{F}' : x^\star_{ij} > 0\}$ and $\bar{y}^\star \leftarrow y^\star$ both supported on $\mathcal{F}'$. For each copy $i' \in \mathcal{F}'$ of $i \in \mathcal{F}$, define its *star cost* as $\sum_{j\in\mathcal{C}':i'\in F_j} f(\frac{2}{2+\delta}d(i,j))$.

We iteratively perform the following procedures. For each $i\in\mathcal{F}$ and $j\in\mathcal{C}'$ such that $x^\star_{ij} > 0$, we sort all copies of $i$ in $\mathcal{F}'$ in non-decreasing order of their current star costs, and choose the first several copies such that their $\bar{y}^\star$ values add up to *exactly* $x^\star_{ij}$. If we need to split a facility $i'$ into two copies to make the sum exact, we replace $i'$ with $\{i'_1, i'_2\}$ in $\mathcal{F}'$, set $\bar{y}^\star_{i'_1}$ to whichever value is needed and $\bar{y}^\star_{i'_2} \leftarrow \bar{y}^\star_{i'} - \bar{y}^\star_{i'_1}$. Remove from $F_j$ all copies of $i$, and add the selected copies to $F_j$ again. For any other $j' \neq j$, if some $i' \in F_{j'}$ is split in two, $F_{j'} \leftarrow F_{j'} \setminus \{i'\} \cup \{i'_1, i'_2\}$.

After the procedures, we set $\mathcal{F} \leftarrow \mathcal{F}'$ and the corresponding $y^\star \leftarrow \bar{y}^\star$, $\{F_j\}_{j\in\mathcal{C}'}$ such that they are supported on $\mathcal{F}$. 1 to 4 are easy to verify, since the original solution to S-LP$(f_{\lambda_1})$ is preserved up to facility duplication. To see 5, consider each (original) facility $i$ and all clients $j$ such that $x^\star_{ij} > 0$, denoted by $J_i \subseteq \mathcal{C}'$. It is easy to see each copy of $i$ only appears in $\bigcup_{j\in J_i} F_j$. We use induction to show that, after each iteration, the difference between the maximum and minimum star costs among all copies of $i$ is at most $\rho U$.

The copies of $i$ and their star costs may only change after an iteration where $i$ is selected. Suppose $J_i = \{j_1, \ldots, j_\ell\}$ and we consider the iterations in the order of $(i, j_1), \ldots, (i, j_\ell)$. As the base case, before $(i, j_1)$ is considered, the claim is true because $i$ has only one copy in $\mathcal{F}'$.

Suppose the claim is true after $(i, j_{t-1})$, $t \geq 1$. In the start of the iteration on $(i, j_t)$, we sort the copies of $i$ in non-decreasing order of their current star costs; each client $j_s, s \geq t$ contributes equally to the star cost of each copy of $i$, including $j_t$ in particular, and the difference between the maximum and minimum is at most $\rho U$, using the induction hypothesis. During this iteration, we remove the contributions of $j_t$ to all copies, and add them back to copies that have the smallest star costs. Since $f(\frac{2}{2+\delta} d(i, j_t)) \leq \rho U$ by (S-LP.7), it is easy to verify that the difference between the maximum and minimum after the iteration is still at most $\rho U$. This finishes the induction.

For facility $i$, we let $\mathcal{F}(i) \subseteq \mathcal{F}'$ be the copies of $i$ after the procedures. It follows that

$$\sum_{i'\in\mathcal{F}(i)} \bar{y}^\star_{i'} \sum_{j\in\mathcal{C}':i'\in F_j} f\left(\frac{2}{2+\delta} d(i,j)\right) = \sum_{j\in J_i} f\left(\frac{2}{2+\delta} d(i,j)\right) \sum_{i'\in\mathcal{F}(i)\cap F_j} \bar{y}^\star_{i'}$$
$$= \sum_{j\in J_i} x^\star_{ij} f\left(\frac{2}{2+\delta} d(i,j)\right) \leq \rho U y^\star_i,$$

where the last inequality is due to (S-LP.8). Hence, the minimum star cost is at most $\rho U y^\star_i / \sum_{i'\in\mathcal{F}(i)} \bar{y}^\star_{i'} = \rho U$, and the maximum is at most $2\rho U$, yielding 5. ◀

### 3.1.1.3 Iterative Rounding

We obtain $y^\star \in [0,1]^{\mathcal{F}}$ and $\{F_j\}_{j\in\mathcal{C}'}$ using Lemma 10. To optimize our approximation factor, we use the following *deterministic* metric discretization. Fix $\tau > 1$; define $D_{-2} = -1, D_{-1} = 0$ and $D_l = \tau^l$ for each $l \geq 0$; let $d'(i,j) = \min\{D_l \geq d(i,j) : l \geq -2\}$. For each $j \in \mathcal{C}'$, we call $F_j$ its *outer ball*, define its *radius level* $l_j \in \mathbb{Z}$ such that $D_{l_j} = \max_{i\in F_j} d'(i,j)$, and define its *inner ball* $B_j = \{i \in F_j : d'(i,j) \leq D_{l_j - 1}\}$. For $0 < \lambda_2 \leq 1/\tau$, we define an auxiliary LP.

$$\min \quad \sum_{j\in C_{\text{part}}} \sum_{i\in F_j} y_i f_{\lambda_2}(d'(i,j)) + \sum_{j\in C_{\text{full}}} \left( \sum_{i\in B_j} y_i f_{\lambda_2}(d'(i,j)) + (1 - y(B_j)) f_{\lambda_2}(D_{l_j}) \right) \tag{A-LP($f_{\lambda_2}$)}$$

$$\begin{aligned}
\text{s.t.} \quad & y(F_j) = 1 && \forall j \in C_{\text{core}} && \text{(A-LP.1)}\\
& 0 \leq y_i \leq 1 && \forall i \in \mathcal{F} && \text{(A-LP.2)}\\
& y(B_j) \leq 1 && \forall j \in C_{\text{full}} && \text{(A-LP.3)}\\
& y(F_j) \leq 1 && \forall j \in C_{\text{part}} && \text{(A-LP.4)}\\
& y(\mathcal{F}) \leq k && && \text{(A-LP.5)}\\
& |C_{\text{full}}| + \sum_{j\in C_{\text{part}}} y(F_j) \geq m'. && && \text{(A-LP.6)}
\end{aligned}$$

The objective of A-LP$(f_{\lambda_2})$ is determined by three subsets of clients $C_{\text{full}}, C_{\text{part}}$, and $C_{\text{core}}$, such that $C_{\text{full}} \cup C_{\text{part}} = \mathcal{C}'$; each client in $C_{\text{full}}$ is to be assigned an open facility relatively close to it, and $C_{\text{core}}$ is used for placing these facilities. Initially, we set $C_{\text{full}} = \emptyset$,

$C_{\text{part}} = \mathcal{C}'$ and $C_{\text{core}} = S_0$; each $i \in S_0$ is called a *virtual client* and its initial radius level is $-1$, since $\sum_{i' \text{ co-located with } i} y^\star_i = 1$ by 3 and $D_{-1} = 0$. We use the following Algorithm 1 to iteratively change $y^\star$ and A-LP($f_{\lambda_2}$).

**Algorithm 1** Iterative Rounding [23].

**Input :** outer balls $\{F_j : j \in \mathcal{C}'\}$, radius levels $\{l_j : j \in \mathcal{C}'\}$, inner balls $\{B_j : j \in \mathcal{C}'\}$, $S_0$

**Output :** an output solution $y'$

$C_{\text{full}} \leftarrow \emptyset, C_{\text{part}} \leftarrow \mathcal{C}', C_{\text{core}} \leftarrow S_0$

**while** *true* **do**

  find an optimal basic feasible solution $y'$ to A-LP($f_{\lambda_2}$)

  **if** *there exists $j \in C_{part}$ such that $y'(F_j) = 1$* **then**

    $C_{\text{part}} \leftarrow C_{\text{part}} \setminus \{j\}, C_{\text{full}} \leftarrow C_{\text{full}} \cup \{j\}, B_j \leftarrow \{i \in F_j : d'(i,j) \le D_{l_j - 1}\}$, **update-**$C_{\text{core}}(j)$

  **else if** *there exists $j \in C_{full}$ such that $y'(B_j) = 1$* **then**

    $l_j \leftarrow l_j - 1, F_j \leftarrow B_j, B_j \leftarrow \{i \in F_j : d'(i,j) \le D_{l_j - 1}\}$, **update-**$C_{\text{core}}(j)$

  **else** break

**return** $y'$

**update-**$C_{\textbf{core}}(j)$

  **if** *there exists no $j' \in C_{core}$ with $l_{j'} \le l_j$ and $F_j \cap F'_j \neq \emptyset$* **then**

    remove from $C_{\text{core}}$ all $j'$ such that $F_j \cap F_{j'} \neq \emptyset$, $C_{\text{core}} \leftarrow C_{\text{core}} \cup \{j\}$

▶ **Lemma 11.** *In each iteration, $y'$ is feasible after modifying the LP. The objective value of $y'$ is non-increasing throughout the algorithm.*

**Proof.** There are two cases. The first is when we move some $j$ from $C_{\text{part}}$ to $C_{\text{full}}$ when $y'(F_j) = 1$. Since $B_j \subseteq F_j$, it satisfies the new constraints in (A-LP.3) and (A-LP.1), if it is added to $C_{\text{core}}$. Since $F_j = B_j \cup (F_j \setminus B_j)$, the contribution of $j$ to the new objective is the same as when it is in $C_{\text{part}}$, because each $i \in F_j \setminus B_j$ satisfies $d'(i,j) = D_{l_j}$ by definition.

The second case is when we decrease the radius level $l_j$ and invoke the subroutine on $j \in C_{\text{full}}$, where $y'(B_j) = 1$. Comparing the contributions of $j$ before and after the iteration, they are equal since the old contribution has $1 - y'(B_j) = 0$, and we can partition the new outer ball $F_j \leftarrow B_j$ in the same way as above.

In both cases, the objective of $y'$ does not change during an iteration. At the beginning of each iteration, we solve for an optimal basic feasible solution, thus the lemma follows. ◀

▶ **Property 12.** *After each iteration of Algorithm 1, the following properties hold.*

**(12.1)** *$C_{full}$ and $C_{part}$ form a partition of $\mathcal{C}'$, $S_0 \subseteq C_{core}$ and $C_{core} \setminus S_0 \subseteq C_{full}$.*
**(12.2)** *$\{F_j : j \in C_{core}\}$ are mutually disjoint.*
**(12.3)** *For each $j \in \mathcal{C}'$, $D_{l_j} \le \tau R_j$.*
**(12.4)** *For each $j \in \mathcal{C}'$, $l_j \ge -1$.*
**(12.5)** *For each $i$ not co-located with $S_0$, $\sum_{j \in \mathcal{C}' : i \in F_j} f(\frac{2}{2+\delta} d(i,j)) \le 2\rho U$.*

**Proof.** First, by iteratively decreasing the radius levels, we claim that no client can have a radius level of $-2$ or smaller. This is because when $l_j = -1$, its inner ball is $B_j = \{i \in F_j : d'(i,j) \le D_{-2} = -1\} = \emptyset$, the constraint $y(B_j) \le 1$ cannot be tight, and we will not invoke the subroutine **update-**$C_{\text{core}}$ on $j$. This shows (12.4).

To see (12.1), we only need to show that virtual clients in $S_0$ cannot be removed from $C_{\text{core}}$. From the subroutine **update-**$C_{\text{core}}$, $j'$ can be removed by $j$ only when $l_j < l_{j'}$. But each virtual client starts with a radius level of $-1$, and removing any such virtual client means a radius level of $-2$, a contradiction.

(12.2) clearly follows by the definition of the subroutine. (12.3) is due to $F_j \subseteq \mathsf{Ball}_{\mathcal{F}}(j, R_j)$ at the beginning of iterative rounding, hence $D_{l_j} \le \max_{i\in F_j} d'(i,j) \le \tau \max_{i\in F_j} d(i,j) \le \tau R_j$. Lastly, since each $F_j$, $j \in \mathcal{C}'$ is inclusion-wise non-increasing during Algorithm 1, the sum in (12.5), being the star cost of $i$, is also non-increasing and at most $2\rho U$ due to 5. This yields (12.5). ◀

We now establish the connection between S-LP($f_{\lambda_1}$) and A-LP($f_{\lambda_2}$), making it possible to compare their objectives, before and after the iterative rounding process.

▶ **Lemma 13.** *For each $\lambda_1 \in (0, 1]$ and $\lambda_2 = \lambda_1/\tau$, the solution $y^\star$ obtained in Lemma 10 is feasible to* A-LP($f_{\lambda_2}$) *with its objective not increased.*

**Proof.** $y^\star$ is the same as the optimal solution for S-LP($f_{\lambda_1}$) up to facility duplication. Before Algorithm 1, we have $C_{\text{core}} = S_0$. Because we require $y_i = 1$ for each $i \in S_0$ in S-LP($f_{\lambda_1}$) and we can let $F_i$ consist of all copies of $i$ (as a virtual client), (A-LP.1) is satisfied by $y^\star$. Initially, $C_{\text{full}}$ is empty and we only have $y(F_j) \le 1$ for $j \in \mathcal{C}'$ (i.e., (A-LP.4)) and $\sum_{j\in\mathcal{C}'} y(F_j) \ge m'$ (i.e., (A-LP.6)), which are also satisfied by $y^\star$ due to Lemma 10. Therefore, $y^\star$ is indeed feasible to A-LP($f_{\lambda_2}$).

In S-LP($f_{\lambda_1}$), each facility-client pair $(i, j)$ has a contribution of $x^\star_{ij} f_{\lambda_1}(d(i,j))$. In A-LP($f_{\lambda_2}$), because $C_{\text{full}} = \emptyset$, its contribution is $y^\star_i f_{\lambda_2}(d'(i,j))$ only when $i \in F_j$ and zero otherwise. Since $d'$ is rounded-up by a factor of at most $\tau$ and $\lambda_1 = \tau\lambda_2$, we further obtain

$$y^\star_i f(\lambda_2 d'(i,j)) \le y^\star_i f(\lambda_1 d(i,j)),$$

thus for $y^\star$, the objective of A-LP($f_{\lambda_2}$) is at most the objective of S-LP($f_{\lambda_1}$). ◀

▶ **Lemma 14.** *There are at most two fractional variables in the output solution $y'$. At the conclusion of the algorithm, for each $j \in C_{full}$,*

$$\sum_{i\in\mathcal{F}:d(i,j)\le\frac{3\tau-1}{\tau-1}D_{l_j}} y'_i \ge 1.$$

**Proof.** Since $y'$ is an optimal basic feasible solution, if it has $t > 0$ strictly fractional variables, there are at least $t$ non-trivial (i.e., not in the form of $y_i \ge 0$ or $y_i \le 1$) and independent constraints of A-LP($f_{\lambda_2}$) that are tight at $y'$ (see, e.g., [25]). The remaining constraints form a knapsack constraint (A-LP.6), and a laminar family (A-LP.1) plus (A-LP.5), according to (12.2). The number of such tight independent constraints is therefore at most $t/2 + 1$. This means that $t/2 + 1 \ge t$, and thus $t \in \{1, 2\}$.

To show the second assertion, we first use induction to show that, for each $j$ that is added to $C_{\text{core}}$ during Algorithm 1 with radius level $l$, the final solution satisfies $\sum_{i\in\mathcal{F}:d(i,j)\le\frac{\tau+1}{\tau-1}D_l} y'_i \ge 1$. The base case is simple for $l = -1$, since we know such $j$ cannot be removed from $C_{\text{core}}$, and $y'$ satisfies the inequality due to (A-LP.1). Suppose the claim is true up to $l - 1$, $l \ge 0$. For $j$ added to $C_{\text{core}}$ with radius level $l$, if it is not later removed from $C_{\text{core}}$, the claim directly follows from (A-LP.1). Otherwise, if $j$ is later removed by $j'$ with $l_{j'} < l_j = l$ and $F_{j'} \cap F_j \ne \emptyset$, using the induction hypothesis, the inequality holds for $j'$ and $l_{j'}$, where $D_{l_{j'}} \le D_l/\tau$. Using the triangle inequality, all these facilities are at a distance at most $\frac{\tau+1}{\tau-1}D_{l_{j'}} + D_{l_{j'}} + D_l \le (\frac{\tau+1}{\tau(\tau-1)} + \frac{1}{\tau} + 1)D_l = \frac{\tau+1}{\tau-1}D_l$ from $j$, showing the induction step.

Back to the proof of the lemma. When we invoke $j$ on its *final* radius $l_j$, if we can indeed add $j$ to $C_{\text{core}}$, the claim above is sufficient since $\frac{\tau+1}{\tau-1} \le \frac{3\tau-1}{\tau-1}$. If it cannot be added to $C_{\text{core}}$, it is because there exists $j' \in C_{\text{core}}$ with $F_{j'} \cap F_j \neq \emptyset$ and $l_{j'} \le l_j$. Using the claim on the iteration when we add $j'$ to $C_{\text{core}}$ with radius level $l_{j'}$, and using the triangle inequality, all the facilities in the sum are at a distance at most $\frac{\tau+1}{\tau-1} D_{l_{j'}} + D_{l_{j'}} + D_{l_j} \le \frac{3\tau-1}{\tau-1} D_{l_j}$, whence the lemma follows. ◀

The following lemma is concerned with the objective value of $y'$ in $\mathrm{LP}(f_\lambda)$, where we replace $(m, \mathcal{C})$ with $(m', \mathcal{C}')$ and use the name $\mathrm{LP}(f_\lambda)$ here with a slight abuse of notation.

▶ **Lemma 15.** *Let $0 < \lambda \le \frac{2\tau-2}{\tau(3\tau-1)(2+\delta)}$. Let $y^\star$ and outer balls $\{F_j\}_{j\in\mathcal{C}'}$ be obtained from Lemma 10 and $\text{S-LP}(f_{\lambda_1})$ where $\lambda_1 = \frac{\tau(3\tau-1)}{\tau-1}\lambda$. The iterative rounding algorithm on $\text{A-LP}(f_{\lambda_2})$ where $\lambda_2 = \frac{3\tau-1}{\tau-1}\lambda$ returns a solution $y'$ with at most two fractions. Moreover, $y'$ is a feasible solution to $\mathrm{LP}(f_\lambda)$ with objective at most $\frac{\lambda\tau(3\tau-1)(2+\delta)}{2\tau-2}U'$.*

**Proof.** From Lemma 9 and Lemma 13, $y^\star$ is feasible to $\text{A-LP}(f_{\lambda_2})$ and its objective value is upper bounded by $\frac{\lambda_1(2+\delta)}{2}U' = \frac{\lambda\tau(3\tau-1)(2+\delta)}{2(\tau-1)}U'$. From Lemma 14, the final solution $y'$ has at most two fractional values; if we further take $y'$ to $\mathrm{LP}(f_\lambda)$, we can assign each client in $C_{\text{full}} \setminus C_{\text{core}}$ to an extent of 1 to facilities at most $\frac{3\tau-1}{\tau-1}D_{l_j}$ away; the feasibility of $y'$ w.r.t. $\mathrm{LP}(f_\lambda)$ is guaranteed by Lemma 11 and Property 12. From Lemma 11 and Lemma 13, the objective value of $y'$ in $\mathrm{LP}(f_\lambda)$ is upper-bounded by (recall that $d' \ge d$ and $C_{\text{full}} \cup C_{\text{part}}$ is a partition of $\mathcal{C}'$)

$$\sum_{j\in C_{\text{part}}}\sum_{i\in F_j} y'_i f_\lambda(d'(i,j)) + \sum_{j\in C_{\text{full}}}\left(\sum_{i\in B_j} y'_i f_\lambda(d'(i,j)) + (1-y'(B_j)) f_\lambda\left(\frac{3\tau-1}{\tau-1}D_{l_j}\right)\right). \quad (2)$$

Because $\lambda_2 = \lambda\frac{3\tau-1}{\tau-1}$, (2) is at most the objective of $\text{A-LP}(f_{\lambda_2})$ and $\le \frac{\lambda\tau(3\tau-1)(2+\delta)}{2\tau-2}U'$. ◀

The following theorem converts the almost-integral solution $y'$ to an integral one $\hat{y}$.

▶ **Theorem 16.** *There exists $\lambda > 0$ depending on $\delta$ and $\tau$ such that we can efficiently compute an integral solution $\hat{y}$ to $\mathrm{LP}(f_\lambda)$ with objective value at most $5\rho U$ larger than that of $y'$.*

**Proof.** The case of less than 2 fractions are easier thus omitted here; in the rest of the proof, suppose $y'_{i_1}$ and $y'_{i_2}$ are the two fractional variables. Because $y'$ is a basic feasible solution, we must have $y'_{i_1} + y'_{i_2} = 1$ since the tight constraints in $\text{A-LP}(f_{\lambda_2})$ represent the intersection of a laminar family and a knapsack constraint. Let $C_1 = \{j \in C_{\text{part}} : i_1 \in F_j, i_2 \notin F_j\}$ and $C_2 = \{j \in C_{\text{part}} : i_1 \notin F_j, i_2 \in F_j\}$. W.l.o.g., we assume $|C_1| \ge |C_2|$. One has $|C_1| + |C_{\text{full}}| \ge y'_{i_1}|C_1| + y'_{i_2}|C_2| + |C_{\text{full}}| \ge m'$ using (A-LP.6). Define $\hat{y}$ such that $\hat{y}_{i_1} = 1$, $\hat{y}_{i_2} = 0$ and $\hat{y}_i = y'_i$ for $i \in \mathcal{F} \setminus \{i_1, i_2\}$. Let $\hat{F} = \{i \in \mathcal{F} : \hat{y}_i = 1\}$ and $\hat{C}' = C_1 \cup C_{\text{full}}$.

Using (12.5), the extra cost of assigning all of $C_1$ to $i_1$ is at most

$$\sum_{j\in C_1} f_\lambda(d(i_1,j)) \le \sum_{j\in C_1} f\left(\frac{2d(i_1,j)}{2+\delta}\right) \le \sum_{j:i_1\in F_j} f\left(\frac{2d(i_1,j)}{2+\delta}\right) \le 2\rho U \Leftarrow \lambda \le \frac{2}{2+\delta}. \quad (3)$$

Next, because we reduce the extent of opening $i_2$ to zero, it remains to bound the extra cost of re-assigning full clients that were assigned to $i_2$, defined as $J = \{j \in C_{\text{full}} : i_2 \in B_j\}$; we choose $J$ here because these are the *full* clients whose contributions are changed in (2). Let $\gamma > 0$ be some constant which we will determine later. Let $i^\star$ be the nearest facility to $i_2$ in $\hat{F}$ and $t' = d(i_2, i^\star)$. Let $J_1 = \{j \in J : d(j, i_2) > \gamma t'\}$ and $J_2 = \{j \in J : d(j, i_2) \le \gamma t'\}$. For $j \in J_1$, we have $d(j, i^\star) \le d(j, i_2) + d(i_2, i^\star) < (1 + \frac{1}{\gamma})d(j, i_2)$, thus if we want the following upper bound on the extra assignment costs,

$$\sum_{j \in J_1} f(\lambda d(j, i^\star)) \le \sum_{j \in J_1} f\left(\frac{2}{2+\delta} d(j, i_2)\right) \le 2\rho U \Leftarrow \lambda \le \frac{2}{2+\delta} \cdot \frac{\gamma}{1+\gamma}. \tag{4}$$

For $j \in J_2$, let $i'$ be the nearest open facility to $j$. It is easy to verify that $d(j, i') \le \frac{3\tau-1}{\tau-1} D_{l_j}$ using Lemma 14 and the definition of $\hat{y}$. Since $i^\star$ is the nearest to $i_2$, one has

$$d(j, i') \ge d(i_2, i') - d(j, i_2) \ge d(i_2, i^\star) - \gamma t' = (1-\gamma)t'$$

using the triangle inequality, and thus $R_j \ge D_{l_j}/\tau \ge \frac{\tau-1}{\tau(3\tau-1)}(1-\gamma)t'$. Let $t = \frac{\tau-1}{\tau(3\tau-1)}(1-\gamma)t'$. Suppose that $\frac{\delta}{4+3\delta} t \ge \gamma t'$, then using (8.2) and recalling $R_j = (1+3\delta/4)\widehat{R}_j$, we have

$$|J_2| \le \left|\left\{ j \in \mathsf{Ball}_{\mathcal{C}'}\left(i_2, \frac{\delta}{4+3\delta} t\right) : R_j \ge t \right\}\right| \le \frac{\rho U}{f(\frac{(1-\delta)(1-\delta/4)}{1+3\delta/4} t)}.$$

Using the triangle inequality, we have $d(j, i^\star) \le (1+\gamma)t'$. The total extra cost of assigning $J_2$ to $i^\star$ is at most

$$\sum_{j \in J_2} f(\lambda d(j, i^\star)) \le f(\lambda(1+\gamma)t')|J_2| \le \rho U \Leftarrow \lambda \le \frac{(1-\delta)(1-\delta/4)}{(1+3\delta/4)} \cdot \frac{\tau-1}{\tau(3\tau-1)} \cdot \frac{1-\gamma}{1+\gamma}. \tag{5}$$

Denote $\sigma = \frac{\tau-1}{\tau(3\tau-1)}$ and let $\gamma = \frac{\delta\sigma}{4+3\delta+\delta\sigma}$ so that $\frac{\delta}{4+3\delta} t = \gamma t'$. By letting $\lambda$ be the minimum of (3)(4)(5) and summing over the three cases, the increase of objective value w.r.t. $\mathrm{LP}(f_\lambda)$ is at most $2\rho U + 2\rho U + \rho U = 5\rho U$, thus the theorem follows. ◀

#### 3.1.1.4 Proof of Theorem 1

Let $\frac{\tau-1}{\tau(3\tau-1)} = 0.101$ and $\delta = 0.81765$, thus $\lambda \in (0.008856, 0.008857)$ (see the proof of Theorem 16). We fix $\epsilon > 0$ and obtain a faithful function $f$ using Theorem 3. Fix two small constants $\delta, \rho > 0$, compute $\mathcal{C}', m', S_0, \{R_j : j \in \mathcal{C}'\}$ via Theorem 7 and Theorem 8, and solve $\text{S-LP}(f_{\lambda_1})$ with $\lambda_1 = \frac{\tau(3\tau-1)}{\tau-1}\lambda$. Using iterative rounding, we obtain an almost-integral solution to $\mathrm{LP}(f_\lambda)$ using Lemma 15. Next, we compute an integral solution $\hat{y}$ using Theorem 16, with the objective w.r.t. $\mathrm{LP}(f_\lambda)$ increased by at most $5\rho U$.

Let $\hat{F} = \{i \in \mathcal{F} : \hat{y}_i = 1\}$ and $\hat{C}' = C_1 \cup C_{\mathrm{full}}$ as in the proof of Theorem 16. There are at least $m'$ clients in $\hat{C}'$. We assume $|\hat{C}'| \le m$, otherwise the following argument is simpler. We greedily connect $m - |\hat{C}'| \le m - m'$ clients in $\mathcal{C} \setminus \mathcal{C}'$ to their nearest open facilities in $\hat{F}$, minimizing $f\left(\frac{1-\delta}{1+\delta} d(j, \hat{F})\right)$ for each of them and output the final solution $(\hat{F}, \hat{C})$. We consider the objective of $(\hat{F}, \hat{C})$ in $\mathscr{J}_{f_\lambda}$ on $\hat{C} \setminus \mathcal{C}'$ and $\hat{C} \cap \mathcal{C}'$ separately.

$$\begin{aligned}
&\sum_{j \in \hat{C} \setminus \mathcal{C}'} f(\lambda d(j, \hat{F})) + \sum_{j \in \hat{C} \cap \mathcal{C}'} f(\lambda d(j, \hat{F})) \\
&\le \sum_{j \in C^\star \setminus \mathcal{C}'} \frac{\lambda(1+\delta)}{1-\delta} f\left(\frac{1-\delta}{1+\delta} d(j, \hat{F})\right) + \left(\frac{\lambda\tau(3\tau-1)(2+\delta)}{2\tau-2} U' + 5\rho U\right) \\
&\le \max\left\{\frac{\lambda(1+\delta)}{1-\delta}, \frac{\lambda\tau(3\tau-1)(2+\delta)}{2\tau-2}\right\} \left(\sum_{j \in C^\star \setminus \mathcal{C}'} f\left(\frac{1-\delta}{1+\delta} d(j, S_0)\right) + U'\right) + 5\rho U \\
&\le 0.12354U + 5\rho U.
\end{aligned} \tag{6}$$

In the above, the first inequality is due to Lemma 15, Theorem 16 and the greedy selection of $\hat{C} \setminus \mathcal{C}'$. The second is because $S_0 \subseteq \hat{F}$. The last is due to (7.4) and our choices of parameters.

Recall that $U \leq (1+\epsilon)V^\star$ where $V^\star$ is the optimal objective of $\mathscr{I}_f$ and $U \leq (1+O(\epsilon))\mathsf{opt}$. Using (6) and Theorem 3 again, the objective of $(\hat{F}, \hat{C})$ in the original instance $\mathscr{I}$ is at most $(0.12354U + 5\rho U + (1+O(\epsilon))\mathsf{opt})/\lambda \leq (126.9 + O(\epsilon+\rho))\mathsf{opt} \leq 127\mathsf{opt}$, by choosing small enough $\rho$ and $\epsilon$. The running time is obtained from the enumeration process and bounded by a polynomial. ◀

## 3.2 Ordered Matroid/Knapsack Median

We consider the ordered matroid median problem (OMatMed) and the ordered knapsack median problem (OKnapMed). Formally, in OMatMed, we instantiate OrdClst such that $\mathscr{F}$ is the set of independent sets of an arbitrary matroid $\mathcal{M} = (\mathcal{F}, \mathscr{F})$ and $\mathscr{C} = \{\mathcal{C}\}$; in OKnapMed, we instantiate OrdClst such that each facility $i \in \mathcal{F}$ has a weight $\mathsf{wt}_i \geq 0$, $\mathscr{F}$ is the set of facility subsets with total weight at most $W$, i.e., $\mathscr{F} = \{F \subseteq \mathcal{F} : \sum_{i \in F} \mathsf{wt}_i \leq W\}$ and $\mathscr{C} = \{\mathcal{C}\}$. It is easy to see that OMatMed and OKnapMed generalize matroid center and matroid median, knapsack center and knapsack median, respectively. Moreover, since the cardinality constraint $|F| \leq k$ is trivially recovered by the matroid and knapsack constraints, O$k$Med is also generalized by OMatMed and OKnapMed.

Theorem 2 is obtained using the same reduction by Theorem 3 and similar iterative rounding algorithms as RO$k$Med. We provide the details of the algorithms and the proofs in Appendix B and Appendix C.

▶ Remark. We remark on the difficulties of OMatMed and OKnapMed under previous methods for O$k$Med. The integrality gap in the natural relaxation for matroid median is a constant (see, e.g., [22]), thus it is likely that the algorithm by Byrka et al. [5] could provide a constant-factor approximation for OMatMed after some modifications; this would hardly be surprising since our reduction algorithm also gives a simpler analysis for OMatMed, compared with RO$k$Med. For OKnapMed, however, it appears that previous methods will fail due to the unbounded integrality gap in the natural relaxation for knapsack median (see, e.g., [24]); our reduction framework can circumvent this issue analogously to RO$k$Med.

# 4 Conclusion

In this paper, we present a reduction framework for a class of clustering problems with ordered objectives, which preserves the approximation guarantee up to constant factors. This leads to the first polynomial-time constant-factor approximation algorithms for three natural clustering problems, namely, robust ordered $k$-median, ordered matroid median and ordered knapsack median. We find the problem of robust ordered $k$-median particularly interesting, since its objective exhibits a certain unimodal shape, which can be nicely motivated by real-world applications.

We list some open questions here that we find interesting.

- Our reduction framework is based upon the sparsification methods proposed by Aouad and Segev [1] and Byrka et al. [5]. On the ordered objective and symmetric monotone norms in general, there have been other approximation methods, e.g., [8, 19]. It would be interesting to see whether our approximation guarantees can be improved by leveraging other techniques and ideas in the literature.
- Although the objective of robust ordered $k$-median is distinctly unimodal in its shape (see Figure 1), there are still more general unimodal objective functions that are not captured by it. Obtaining an approximation algorithm for arbitrary unimodal vectors, even with an $O(\log n)$-factor approximation guarantee, is beyond the scope of our current framework, so it might require some brand new ideas to handle these objectives.

## References


1 Ali Aouad and Danny Segev. The ordered $k$-median problem: surrogate models and approximation algorithms. *Math. Program.*, 177(1-2):55–83, 2019. doi:10.1007/s10107-018-1259-3.

2 Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. Local search heuristics for $k$-median and facility location problems. *SIAM J. Comput.*, 33(3):544–562, 2004. doi:10.1137/S0097539702416402.

3 Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Joachim Spoerhase, Aravind Srinivasan, and Khoa Trinh. An improved approximation algorithm for knapsack median using sparsification. *Algorithmica*, 80(4):1093–1114, 2018. doi:10.1007/s00453-017-0294-4.

4 Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for $k$-median and positive correlation in budgeted optimization. *ACM Trans. Algorithms*, 13(2):23:1–23:31, 2017. doi:10.1145/2981561.

5 Jaroslaw Byrka, Krzysztof Sornat, and Joachim Spoerhase. Constant-factor approximation for ordered $k$-median. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 620–631, 2018. doi:10.1145/3188745.3188930.

6 Deeparnab Chakrabarty, Prachi Goyal, and Ravishankar Krishnaswamy. The non-uniform $k$-center problem. *ACM Trans. Algorithms*, 16(4):46:1–46:19, 2020. doi:10.1145/3392720.

7 Deeparnab Chakrabarty and Chaitanya Swamy. Interpolating between $k$-median and $k$-center: Approximation algorithms for ordered $k$-median. In *45th International Colloquium on Automata, Languages, and Programming*, volume 107 of *LIPIcs*, pages 29:1–29:14, 2018. doi:10.4230/LIPIcs.ICALP.2018.29.

8 Deeparnab Chakrabarty and Chaitanya Swamy. Approximation algorithms for minimum norm and ordered optimization problems. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 126–137, 2019. doi:10.1145/3313276.3316322.

9 Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the $k$-median problem. *Journal of Computer and System Sciences*, 65(1):129–149, 2002. doi:10.1006/jcss.2002.1882.

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

11 Moses Charikar and Shi Li. A dependent LP-rounding approach for the $k$-median problem. In *Automata, Languages, and Programming - 39th International Colloquium, Proceedings, Part I*, volume 7391 of *Lecture Notes in Computer Science*, pages 194–205, 2012. doi:10.1007/978-3-642-31594-7_17.

12 Danny Z. Chen, Jian Li, Hongyu Liang, and Haitao Wang. Matroid and knapsack center problems. *Algorithmica*, 75(1):27–52, 2016. doi:10.1007/s00453-015-0010-1.

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

14 Vincent Cohen-Addad, Anupam Gupta, Lunjia Hu, Hoon Oh, and David Saulpic. An improved local search algorithm for $k$-median. In *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms*, pages 1556–1612. SIAM, 2022. doi:10.1137/1.9781611977073.65.

15 Jack R. Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Optimization - Eureka, You Shrink!*, volume 2570 of *Lecture Notes in Computer Science*, pages 11–26. Springer, 2001. doi:10.1007/3-540-36478-1_2.

16 Anupam Gupta, Benjamin Moseley, and Rudy Zhou. Structural iterative rounding for generalized $k$-median problems. In *48th International Colloquium on Automata, Languages, and Programming*, volume 198 of *LIPIcs*, pages 77:1–77:18, 2021. doi:10.4230/LIPIcs.ICALP.2021.77.

17 David G. Harris, Thomas W. Pensyl, Aravind Srinivasan, and Khoa Trinh. A lottery model for center-type problems with outliers. *ACM Transactions on Algorithms*, 15(3):36:1–36:25, 2019. doi:10.1145/3311953.

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

19 Sharat Ibrahimpur and Chaitanya Swamy. Minimum-norm load balancing is (almost) as easy as minimizing makespan. In *48th International Colloquium on Automata, Languages, and Programming*, volume 198 of *LIPIcs*, pages 81:1–81:20, 2021. doi:10.4230/LIPIcs.ICALP.2021.81.

20 Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing*, pages 731–740, 2002. doi:10.1145/509907.510012.

21 Kamal Jain and Vijay V. Vazirani. Approximation algorithms for metric facility location and $k$-median problems using the primal-dual schema and Lagrangian relaxation. *J. ACM*, 48(2):274–296, 2001. doi:10.1145/375827.375845.

22 Ravishankar Krishnaswamy, Amit Kumar, Viswanath Nagarajan, Yogish Sabharwal, and Barna Saha. The matroid median problem. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1117–1130, 2011. doi:10.1137/1.9781611973082.84.

23 Ravishankar Krishnaswamy, Shi Li, and Sai 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. doi:10.1145/3188745.3188882.

24 Amit Kumar. Constant factor approximation algorithm for the knapsack median problem. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 824–832, 2012. doi:10.1137/1.9781611973099.66.

25 Lap Chi Lau, Ramamoorthi Ravi, and Mohit Singh. *Iterative methods in combinatorial optimization*, volume 46. Cambridge University Press, 2011.

26 Shi Li and Ola Svensson. Approximating $k$-median via pseudo-approximation. *SIAM J. Comput.*, 45(2):530–547, 2016. doi:10.1137/130938645.

27 Chaitanya Swamy. Improved approximation algorithms for matroid and knapsack median problems and applications. *ACM Trans. Algorithms*, 12(4):49:1–49:22, 2016. doi:10.1145/2963170.


## A Missing Proofs for Robust Ordered $k$-Median

**Proof of Theorem 7.** Let us first assume we know $(F^\star, C^\star)$; we will remedy this assumption after the construction of $\mathscr{J}'$. Recall that $U \in [V^\star, (1+\epsilon)V^\star)$ and $V^\star = \sum_{j\in C^\star} f(c_j^\star)$.

Set $S_0 = \emptyset$, $\mathcal{C}' = \mathcal{C}$. Whenever there exists $i \in F^\star \setminus S_0$ such that $\sum_{j\in C^\star:\kappa_j^\star = i} f(c_j^\star) \geq \rho U$, we set $S_0 \leftarrow S_0 \cup \{i\}$. This process can be repeated for at most $O(1/\rho)$ times, because the subsets of clients assigned to the facilities in $S_0$ are mutually disjoint by the definition of $\kappa_j^\star$, and the overall sum of $f$ values is at most $V^\star \leq U$. The remaining facilities in $F^\star \setminus S_0$ will always satisfy (7.2) because we will only add facilities to $S_0$ and remove clients from $\mathcal{C}'$.

Next, we put $C'^\star = C^\star \cap \mathcal{C}'$ at all times. Whenever there exists $p \in \mathcal{F} \cup \mathcal{C}'$ such that $\left|\mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)\right| \cdot f((1-\delta)c_p^\star) \geq \rho U$, set $\mathcal{C}' \leftarrow \mathcal{C}' \setminus \mathsf{Ball}_{\mathcal{C}'}\left(p, \delta c_p^\star\right)$ and $S_0 \leftarrow S_0 \cup \{\kappa_p^\star\}$. Each removed client $j$ is from $\mathcal{C}'$ and satisfies $f(c_j^\star) \geq f(c_p^\star - d(j,p)) \geq f((1-\delta)c_p^\star)$ using the triangle inequality. Thus the total $f$ value removed is at least $\left|\mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)\right| \cdot f((1-\delta)c_p^\star) \geq \rho U$. Using a similar argument, this process can also be repeated for at most $O(1/\rho)$ times. The condition in (7.3) is then satisfied by definition.

Note that each such removed client $j \in \mathcal{C} \setminus \mathcal{C}'$ has, by the triangle inequality again,

$$f\left(\frac{1-\delta}{1+\delta} d(j, S_0)\right) \leq f\left(\frac{1-\delta}{1+\delta}(d(j,p) + d(p, \kappa_p^\star))\right) \leq f((1-\delta)c_p^\star) \leq f(c_j^\star),$$

where the last inequality is because $c_j^\star \ge c_p^\star - d(j,p) \ge (1-\delta)c_p^\star$. Therefore, by summing over all $j \in C^\star$, (7.4) follows since

$$\sum_{j\in C^\star\setminus\mathcal{C}'} f\left(\frac{1-\delta}{1+\delta}d(j,S_0)\right) + U' \le \sum_{j\in C^\star\setminus\mathcal{C}'} f(c_j^\star) + \sum_{j\in C^\star\cap\mathcal{C}'} f(c_j^\star) = V^\star \le U.$$

Finally, we remove the dependence of the procedures on $(F^\star, C^\star)$, by noticing that $|S_0| = O(1/\rho)$, and $\mathcal{C}'$ is obtained from $\mathcal{C}$ by removing $O(1/\rho)$ closed balls. Since $m' = |C^\star\cap\mathcal{C}'|$ only takes values in $[m]$, the total number of possible outcomes is at most $n_0^{O(1/\rho)}$, and we can simply enumerate all possible configurations of $(\mathcal{C}', m', S_0)$. (7.1) also follows by definition. ◀

**Proof of Theorem 8.** We iteratively construct $\{\widehat{R}_j : j \in \mathcal{C}'\}$ that always maintain (8.2), then prove (8.1). Initially, let $\widehat{R}_j = 0$ for each $j \in \mathcal{C}'$. In each iteration $k \ge 1$, we try to assign the $k$-th largest distance $t'$ in $\{d(i,j) : i \in \mathcal{F}, j \in \mathcal{C}'\} \setminus \{0\}$ *sequentially* to unassigned clients $\{j \in \mathcal{C}' : \widehat{R}_j = 0\}$ without violating (8.2); it is easy to verify that (8.2) is always maintained, since it suffices to consider the case of $t = t'$ for each $p \in \mathcal{F}\cup\mathcal{C}'$ (cf. [23]).

Recall that $C'^\star = C^\star \cap \mathcal{C}'$. We construct a one-to-one mapping $\phi : C'^\star \to \mathcal{C}'$ and show the solution $(F^\star, \phi(C'^\star))$ satisfies (8.1). Initially, we let $\phi$ be the identity function on $C'^\star$. Consider the clients in $\{j \in C'^\star : c_j^\star > (1+3\delta/4)\widehat{R}_j\}$ in non-decreasing order of $c_j^\star$. For each such $j$, we want to update $\phi(j)$ to an "unused" client in the current $\mathcal{C}' \setminus \phi(C'^\star)$ such that $d(\phi(j), j) \le \delta c_j^\star/2$ and $\widehat{R}_{\phi(j)} \ge c_j^\star$. If such $\phi(j)$ exists for each $j \in C'^\star$, we assign $\phi(j)$ to $\kappa_j^\star$, define $\kappa'_{\phi(j)} = \kappa_j^\star$ and thus $c'_{\phi(j)} = d(\phi(j), \kappa_j^\star) \le c_j^\star + d(j,\phi(j)) \le (1+\delta/2)c_j^\star \le (1+\delta/2)\widehat{R}_{\phi(j)}$. Moreover, one has

$$\forall i \in F^\star\setminus S_0,\quad \sum_{j\in\phi(C'^\star):\kappa'_j=i} f\left(\frac{2}{2+\delta}c'_j\right) \le \sum_{j\in C'^\star:\kappa_j^\star=i} f(c_j^\star) \le \rho U,$$

where the last inequality is due to (7.2). Similarly, one has

$$\sum_{j\in\phi(C'^\star)} f\left(\frac{2}{2+\delta}c'_j\right) = \sum_{j\in C'^\star} f\left(\frac{2}{2+\delta}c'_{\phi(j)}\right) \le \sum_{j\in C'^\star} f(c_j^\star) = U',$$

therefore (8.1) is satisfied by $(F^\star, \phi(C'^\star))$ in this case.

It remains to show that such an unused $j' \in \mathcal{C}'\setminus\phi(C'^\star)$ can always be found for each $j \in C'^\star$ with $c_j^\star > (1+3\delta/4)\widehat{R}_j$. Notice that we have $\widehat{R}_j = 0$ when $t' = c_j^\star$ is considered during the construction, so setting $\widehat{R}_j = c_j^\star$ would be a violation in (8.2); that is, there exists $p \in \mathcal{F}\cup\mathcal{C}'$ such that $d(p,j) \le \delta c_j^\star/4$ and the set $H_j = \left\{j' \in \mathsf{Ball}_{\mathcal{C}'}\left(p, \delta c_j^\star/4\right) : \widehat{R}_{j'} \ge c_j^\star\right\}$ satisfies

$$|H_j \cup \{j\}| = |H_j| + 1 > \frac{\rho U}{f((1-\delta)(1-\delta/4)c_j^\star)}. \tag{7}$$

Further, if there exists some $j' \in H_j \setminus \phi(C'^\star)$, we can set $\phi(j) = j'$ since $\widehat{R}_{j'} \ge c_j^\star$ and $d(j,j') \le d(j',p) + d(p,j) \le \delta c_j^\star/2$ using the triangle inequality. Therefore, it suffices to prove $H_j \not\subseteq \phi(C'^\star)$.

For the sake of contradiction, assume $H_j \subseteq \phi(C'^\star)$ when we want to update $\phi(j)$. For each $\phi(\hat{j}) \in H_j$, we have $d(p,\hat{j}) \le d(p,\phi(\hat{j})) + d(\hat{j},\phi(\hat{j})) \le 3c_j^\star/4$, because we consider the clients in non-decreasing order of $c_j^\star$ and hence $d(\hat{j},\phi(\hat{j})) \le \delta c_{\hat{j}}^\star/2 \le \delta c_j^\star/2$ in earlier iterations. This shows that (currently) $\phi^{-1}(H_j) \subseteq \mathsf{Ball}_{C'^\star}\left(p, 3\delta c_j^\star/4\right)$. We further have $c_p^\star \geq c_j^\star - d(p,j) \geq (1-\delta/4)c_j^\star$ and $\delta c_p^\star \geq \delta(1-\delta/4)c_j^\star \geq (3\delta/4)c_j^\star$ as $\delta < 1$. Thus $\phi^{-1}(H_j) \subseteq \mathsf{Ball}_{C'^\star}\left(p, 3\delta c_j^\star/4\right) \subseteq \mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)$. Because $j \notin \phi^{-1}(H_j)$ and $j \in \mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)$, we have

$$\begin{aligned}\left|\mathsf{Ball}_{C'^\star}\left(p, \delta c_p^\star\right)\right| \cdot f((1-\delta)c_p^\star) &\geq (|H_j|+1) f((1-\delta)c_p^\star) \\ &\geq (|H_j|+1) f((1-\delta)(1-\delta/4)c_j^\star) >_{(7)} \rho U,\end{aligned}$$

which is a contradiction to (7.3). ◀

# B Ordered Matroid Median

In this section, we give the first constant-factor approximation for $\mathsf{OMatMed}$. As is pointed out in [23], the natural LP relaxation for matroid median has a small integrality gap; we skip the pre-processing steps of $\mathsf{RO}k\mathsf{Med}$ and provide a sketch on how the iterative rounding algorithm directly outputs the desired approximate solution. Suppose we have a faithful function $f$ in what follows via Theorem 3 and exhaustive search.

1. Ignore the pre-processing steps (i.e., Theorem 7 and Theorem 8) in the $\mathsf{RO}k\mathsf{Med}$ algorithm. To obtain a natural relaxation for any reduced instance $\mathscr{I}_{f_\lambda}$, replace the cardinality constraint $\sum_{i\in\mathcal{F}} y_i \leq k$ in LP($f_\lambda$) with $\sum_{i\in S} y_i \leq r_{\mathcal{M}}(S)$, $\forall S \subseteq \mathcal{F}$; here, $r_{\mathcal{M}}$ is the rank function of the given matroid $\mathcal{M} = (\mathcal{F}, \mathscr{F})$, and these matroid polytope constraints follow from a classic result by Edmonds [15].
2. We *no longer* have the stronger relaxation as the $\mathsf{RO}k\mathsf{Med}$ case does. We solve the natural relaxation and proceed to the auxiliary LP, which is similar to A-LP($f_{\lambda_2}$) except for the matroid constraints; because each client is fully assigned to an extent of 1 in $\mathsf{OMatMed}$, we also remove (A-LP.6) and change (A-LP.4) to equality constraints. We use Algorithm 1 for iterative rounding; since we do not have $S_0$ or virtual clients, Algorithm 1 starts with $C_{\text{part}} \leftarrow \mathcal{C}$ and $C_{\text{core}} \leftarrow \emptyset$.
3. Because we do not have any outliers and each client will end up in $C_{\text{full}}$, the remaining *tight* constraints after iterative rounding are either from a partition matroid (i.e., $y(F_j) = 1$ for each $j \in C_{\text{core}}$) or from the input matroid (i.e., $\sum_{i\in S} y_i \leq r_{\mathcal{M}}(S)$ for each $S \subseteq \mathcal{F}$). Therefore, the corresponding output solution $y'$ is integral [15].
4. Using the same argument as Lemma 15, the objective value of $y'$ in the natural relaxation LP($f_\lambda$) (we use the same names as $\mathsf{RO}k\mathsf{Med}$ here with a slight abuse of notation) is bounded by the objective of $y'$ in A-LP($f_{\lambda_2}$) where $\lambda_2 = \frac{3\tau-1}{\tau-1}\lambda$, and further bounded by the *optimal* objective of LP($f_{\lambda_1}$) where $\lambda_1 = \frac{\tau(3\tau-1)}{\tau-1}\lambda$. Similar to Lemma 9, the optimum of LP($f_{\lambda_1}$) is at most $\lambda_1(1+O(\epsilon))\mathsf{opt}$, so the objective of $y'$ in LP($f_\lambda$) is also at most $\lambda_1(1+O(\epsilon))\mathsf{opt}$.
5. Using Theorem 3 on the reduced instance $\mathscr{I}_{f_\lambda}$, the integral solution induced by $y'$ has an approximation ratio of

$$\frac{1}{\lambda}\left(\lambda_1(1+O(\epsilon)) + (1+O(\epsilon))\right) = (1+O(\epsilon))\left(\frac{\tau(3\tau-1)}{\tau-1} + \frac{1}{\lambda}\right),$$

where we have $\lambda \leq \frac{\tau-1}{\tau(3\tau-1)} \in (0, 5-2\sqrt{6}]$ because $\lambda_1 \leq 1$ in LP($f_{\lambda_1}$). Therefore, the approximation ratio is minimized when $\lambda = \frac{\tau-1}{\tau(3\tau-1)}$ and $\frac{\tau(3\tau-1)}{\tau-1}$ is minimized, giving

$$10 + 4\sqrt{6} + O(\epsilon) \leq 19.798 + O(\epsilon) \leq 19.8,$$

where we choose a small enough constant $\epsilon > 0$.

# C Ordered Knapsack Median

In this section, we give the first constant-factor approximation for OKnapMed. We closely follow the procedures in [23] and use an iterative rounding algorithm akin to RO$k$Med. Suppose we have a faithful function $f$ in what follows via Theorem 3 and exhaustive search. The following two theorems are similar to Theorem 7 and Theorem 8 in RO$k$Med. Let $(F^\star, C^\star = \mathcal{C})$ be the optimal solution to the original OKnapMed instance $\mathscr{I}$ with optimum $\mathsf{opt} \geq 0$. Recall that we guess $U \in [V^\star, (1+\epsilon)V^\star)$ via binary search, where $V^\star$ is the optimum of $\mathscr{J} = (\mathcal{F}, \mathcal{C}, d, \mathscr{F}, \mathscr{C}, f)$; we have $V^\star \leq (1 + O(\epsilon))\mathsf{opt}$ using Theorem 3.

▶ **Theorem 17.** *Given $\rho, \delta \in (0,1)$ and $U$, there exists an $n_0^{O(1/\rho)}$-time algorithm that finds an* extended *instance $\mathscr{J}' = (\mathcal{F}, \mathcal{C}', d, \mathscr{F}, \mathscr{C}', f, S_0)$ satisfying the following.*

**(17.1)** *$\mathcal{C}' \subseteq \mathcal{C}$, $\mathscr{C}' = \{\mathcal{C}'\}$ and $S_0 \subseteq F^\star$ with $|S_0| = O(1/\rho)$.*

**(17.2)** *For each $i \in F^\star \setminus S_0$, we have $\sum_{j \in \mathcal{C}' : \kappa_j^\star = i} f(c_j^\star) \leq \rho U$,*

**(17.3)** *For each $p \in \mathcal{F} \cup \mathcal{C}'$, we have $\left|\mathsf{Ball}_{\mathcal{C}'}\left(p, \delta c_p^\star\right)\right| \cdot f((1-\delta)c_p^\star) < \rho U$,*

**(17.4)** *Denote $U' = \sum_{j \in \mathcal{C}'} f(c_j^\star)$. We have $\sum_{j \in \mathcal{C} \setminus \mathcal{C}'} f\left(\frac{1-\delta}{1+\delta} d(j, S_0)\right) + U' \leq U$.*

▶ **Theorem 18.** *Given the instance found in Theorem 17, we can efficiently compute a set of upper bounds $\{R_j \geq 0 : j \in \mathcal{C}'\}$ such that for each $j \in \mathcal{C}'$, we have*

$$c_j^\star \leq R_j = \max\left\{R > 0 : \left|\mathsf{Ball}_{\mathcal{C}'}\left(j, \delta R\right)\right| \cdot f((1-\delta)R) \leq \rho U\right\}.$$

The two theorems above are almost identical to those for RO$k$Med, thus we omit their proofs here. By replacing the cardinality constraint $y(F) \leq k$ with the relaxed knapsack constraint $\sum_{i \in \mathcal{F}} \mathsf{wt}_i \cdot y_i \leq W$, and removing the coverage constraint for outliers, we consider a stronger LP similar to S-LP$(f_{\lambda_1})$. We also use iterative rounding on an auxiliary LP similar to A-LP$(f_{\lambda_2})$. Using a similar argument as in Lemma 14, we see that after iterative rounding, the resulting solution $y'$ corresponds to the intersection of a laminar family and a knapsack constraint, hence it contains at most 2 fractional variables. We now focus on obtaining an integral solution $\hat{y}$ from $y'$.

▶ **Theorem 19.** *There exists $\lambda > 0$ depending on $\delta$ and $\tau$, such that we can efficiently compute an integral solution $\hat{y}$ to* LP$(f_\lambda)$ *(in the knapsack case), and its objective value is at most $3\rho U$ larger than that of $y'$.*

**Proof.** If there is only one fractional facility $i_2$, we close it. If there are two, suppose $i_1, i_2$ are the two fractional facilities and $i_1$ is the one with a smaller weight; because $y'$ is a basic feasible solution, we again have $y'_{i_1} + y'_{i_2} = 1$; we fully open $i_1$ and close $i_2$. The set of open facilities $\hat{F}$ is similarly defined as in Theorem 16. Because $\mathsf{wt}_{i_1} \leq \mathsf{wt}_{i_2}$, it is also easy to verify that $\sum_{i \in F} \mathsf{wt}_i \leq W$, thus $\hat{F}$ is indeed a feasible solution.

Unlike RO$k$Med, each client is fully assigned, so it remains to bound the cost incurred from re-assigning clients that were assigned to $i_2$, that is, $J = \{j \in C_{\text{full}} : i_2 \in B_j\}$. Let $\gamma > 0$ be a constant that we determine later, $i^\star$ be the nearest open facility to $i_2$ in $\hat{F}$ and $t' = d(i_2, i^\star)$. Let $J_1 = \{j \in J : d(j, i_2) > \gamma t'\}$ and $J_2 = \{j \in J : d(j, i_2) \leq \gamma t'\}$. For $j \in J_1$, we have $d(j, i^\star) \leq d(j, i_2) + d(i_2, i^\star) < (1 + \frac{1}{\gamma})d(j, i_2)$, thus

$$\sum_{j \in J_1} f(\lambda d(j, i^\star)) \leq \sum_{j \in J_1} f\left(d(j, i_2)\right) \leq 2\rho U \Leftarrow \lambda \leq \frac{\gamma}{1+\gamma}. \tag{8}$$

Fix some $j \in J_2$. Similar as before, we have $R_j \geq D_{l_j}/\tau \geq \frac{\tau-1}{\tau(3\tau-1)}(t' - d(j,i_2)) \geq \frac{\tau-1}{\tau(3\tau-1)}(1-\gamma)t'$. Suppose $\delta R_j \geq 2\gamma t'$, then by Theorem 18 and the triangle inequality,

$$|J_2| \leq |\mathsf{Ball}_{\mathcal{C}'}(i_2, \gamma t')| \leq |\mathsf{Ball}_{\mathcal{C}'}(j, 2\gamma t')| \leq |\mathsf{Ball}_{\mathcal{C}'}(j, \delta R_j)| \leq \frac{\rho U}{f((1-\delta)R_j)}.$$

Using the triangle inequality again, we have $d(j, i^\star) \leq (1+\gamma)t'$ and the following total cost of assigning $J_2$ to $i^\star$ is at most

$$\sum_{j \in J_2} f(\lambda d(j, i^\star)) \leq |J_2| f(\lambda(1+\gamma)t') \leq \rho U \Leftarrow \lambda \leq (1-\delta) \cdot \frac{\tau - 1}{\tau(3\tau-1)} \cdot \frac{1-\gamma}{1+\gamma}. \tag{9}$$

We let $\sigma = \frac{\tau-1}{\tau(3\tau-1)}$ and let $\gamma = \frac{\delta\sigma}{2+\delta\sigma}$ so that $\delta R_j \geq 2\gamma t'$. By letting $\lambda$ be the minimum of (8)(9) and summing over the two cases, the increase of objective value w.r.t. LP$(f_\lambda)$ is at most $2\rho U + \rho U = 3\rho U$, thus the theorem follows. ◀

Let $\delta = 2/3$ and thus $\lambda = \frac{\sigma}{3+2\sigma}$. Similar to Lemma 9, the objective value of S-LP$(f_{\lambda_1})$, $\lambda_1 = \frac{\tau(3\tau-1)}{\tau-1}\lambda$ is at most $\lambda_1 U'$. Using the same argument as Lemma 15, the objective value of $y'$ in the original relaxation LP$(f_\lambda)$ is at most that of A-LP$(f_{\lambda_2})$, $\lambda_2 = \frac{3\tau-1}{\tau-1}\lambda$, which can be bounded by $\lambda_1 U'$ akin to Lemma 13. Using Theorem 19, the objective of $\hat{y}$ to LP$(f_\lambda)$ is at most $\lambda_1 U' + 3\rho U$. Finally, using (17.4) and similarly to (6), the approximation ratio is

$$\left(\max\left\{5\lambda, \frac{\tau(3\tau-1)}{\tau-1}\lambda\right\} + 1 + O(\rho)\right)\frac{1+\epsilon}{\lambda} = \left(\frac{\lambda}{\sigma} + 1 + O(\rho)\right)\frac{1+\epsilon}{\lambda} = \frac{1}{\sigma} + \frac{1}{\lambda} + O(\epsilon + \rho).$$

By letting $\tau = 1 + \sqrt{\frac{2}{3}}$ and $\sigma = 5 - 2\sqrt{6}$, the approximation ratio is at most

$$\frac{4}{\sigma} + 2 + O(\epsilon+\rho) = 22 + 8\sqrt{6} + O(\epsilon+\rho) \leq 41.596 + O(\epsilon+\rho) \leq 41.6,$$

where one chooses $\epsilon$ and $\rho$ that are small enough. The running time is obtained from the enumeration process and bounded by a polynomial.