Interpolating between k-Median and k-Center: Approximation Algorithms for Ordered k-Median

Deepamab Chakrabarty[*] Chaitanya Swamy[†]

Abstract

We consider a generalization of k-median and k-center, called the *ordered k-median* problem. In this problem, we are given a metric space $(\mathcal{D}, \{c_{ij}\})$ with $n = |\mathcal{D}|$ points, and a non-increasing weight vector $w \in \mathbb{R}_+^n$, and the goal is to open k centers and assign each point each point $j \in \mathcal{D}$ to a center so as to minimize $w_1 \cdot$ (largest assignment cost) $+ w_2 \cdot$ (second-largest assignment cost) $+ \ldots + w_n \cdot$ (n-th largest assignment cost). We give an $(18+\epsilon)$-approximation algorithm for this problem. Our algorithms utilize Lagrangian relaxation and the primal-dual schema, combined with an enumeration procedure of Aouad and Segev. For the special case of $\{0,1\}$-weights, which models the problem of minimizing the ℓ largest assignment costs that is interesting in and of by itself, we provide a novel reduction to the (standard) k-median problem, showing that LP-relative guarantees for k-median translate to guarantees for the ordered k-median problem; this yields a nice and clean $(8.5+\epsilon)$-approximation algorithm for $\{0,1\}$ weights.

1 Introduction

We consider the following common generalization of k-median and k-center, which has been referred to as the *ordered k-median* problem [9]. We are given a metric space $(\mathcal{D}, \{c_{ij}\}_{i,j\in\mathcal{D}})$, and an integer $k \geq 0$. We will often refer to points in $\mathcal{D}$ as clients. We are also given non-increasing nonnegative weights $w_1 \geq w_2 \geq \ldots \geq w_n \geq 0$, where $n = |\mathcal{D}|$. For a vector $v \in \mathbb{R}^{\mathcal{D}}$, we use $v^{\downarrow}$ to denote the vector v with coordinates sorted in non-increasing order. That is, we have $v_i^{\downarrow} = v_{\sigma(i)}$, where σ is a permutation of $\mathcal{D}$ such that $v_{\sigma(1)} \geq v_{\sigma(2)} \geq \ldots v_{\sigma(n)}$. The goal in the ordered k-median problem is to choose a set F of k points from $\mathcal{D}$ as centers (or "facilities"), and assign each client $j \in \mathcal{D}$ to a center $i(j) \in F$, so as to minimize

$$cost\big(w; \vec{c} := \{c_{i(j)j}\}_{j\in\mathcal{D}}\big) := w^T \vec{c}^{\downarrow} = \sum_{j=1}^{n} w_j \vec{c}_j^{\downarrow}.$$

Observe that we may assume that, without loss of generality, each client j is assigned to the center $i(j)$ in F that is nearest to it. We may assume that $|\mathcal{D}| > k$, otherwise the problem becomes trivial. It will be useful to notice that, equivalently, we have

$$cost(w; \vec{c}) = \max_{\text{permutations } \pi \text{ of } \mathcal{D}} \sum_{i=1}^{n} w_i \vec{c}_{\pi(i)}$$

which shows that $cost(w; x)$ is a convex function of x, and in fact a seminorm on $\mathbb{R}^{\mathcal{D}}$.

[*] deeparnab@gmail.com. Dept. of Computer Science, Dartmouth College, Hanover, NH 03755-3510, USA.

[†] cswamy@uwaterloo.ca. Dept. of Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1, Canada.
Supported in part by NSERC grant 327620-09 and an NSERC Discovery Accelerator Supplement Award.

Observe that setting $w_1 = 1 = w_2 = \ldots = w_n$ yields the k-median problem; on the other hand, by setting $w_1 = 1,\ w_2 = 0 = w_3 = \ldots = w_n$, we obtain the k-center problem. Thus, the ordered k-median problem nicely interpolates between the k-median and k-center problems. In particular, an interesting case is the setting with $\{0, 1\}$ weights, which means that for some $\ell \in [n]$, we have $w_1 = \ldots = w_\ell = 1$, and $w_{\ell+1} = 0 = \ldots = w_n$; this captures the problem of minimizing the ℓ largest assignment costs, which Tamir [10] calls the *ℓ-centrum problem*.

While the special cases of k-median and k-center have been considered extensively from the viewpoint of developing approximation algorithms, much less is known about the approximability of the ordered k-median problem, especially in general metrics. Aouad and Segev [2] obtained a logarithmic-approximation ratio for general metrics, and Alamdari and Shmoys [1] obtain a bicriteria approximation for the special case, where w is a convex combination of $(1, 0, \ldots, 0)$ and $\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$. For other work related to location theory and ordered-median models, we refer the reader to [8, 7].

In our work, we develop an $(18 + \epsilon)$-approximation algorithm for the ordered k-median problem. In Section 2, we first develop constant-factor approximation algorithms for the case of $\{0, 1\}$-weights, which introduces many of the ideas needed to handle the general setting. In Section 3, we generalize these ideas to obtain constant-factor approximation algorithms for the ordered k-median problem with general weights.

1.1 Relationship with the work of Byrka et al. [3]

Very recently, we learnt that Byrka et al. [3] have also obtained an $O(1)$-approximation guarantee (equal to $38+\epsilon$) for the ordered k-median problem. Our work was done independently and concurrently. In particular, our results for $\{0, 1\}$ weights were obtained independently. We use somewhat different techniques, and obtain an approximation factor that is better than the one obtained in [3] (for $\{0, 1\}$ weights) via a simpler algorithm and analysis.

But we would like to make it clear that it was after we learnt of the work of [3] that we realized that our results can be extended to the general weighted setting. Again, our algorithms and analyses here utilize somewhat different techniques.

2 The setting with $\{0, 1\}$-weights

We first consider the setting with $\{0, 1\}$ weights. Let $\ell \in [n]$ be such $w_1 = \ldots = w_\ell = 1$, $w_{\ell+1} = 0 = \ldots = w_n$. We abbreviate $cost(w; \vec{c})$ to $cost(\ell; \vec{c})$, or simply $cost(\vec{c})$. The $\{0, 1\}$-weight setting serves as a natural starting point for two reasons. First, the problem of minimizing the ℓ most expensive assignment costs is a natural, well-motivated problem that is interesting in its own right. Second, the study of the $\{0, 1\}$-case serves to introduce some of the key underlying ideas that are also used to handle the general setting. Notice also that a non-decreasing weight vector w can be written as a nonnegative linear-combination of such $\{0, 1\}$ weight vectors.

The natural LP-relaxation for this ℓ-centrum problem is to augment the natural LP-relaxation for k-median by introducing a new variable λ to denote the objective value and impose constraints enforcing that the total assignment cost of any set of ℓ clients is at most λ. One can show however that this natural LP has an $\Omega(\ell)$ integrality gap.

Our constant-factor approximation algorithm is based on an alternate novel LP-relaxation of the problem. Our relaxation is based on the following key insight. Suppose there is a solution of objective value B, and we aim to find a solution of objective value $O(B)$. Then, it suffices to find a solution where the total assignment cost of clients having assignment cost larger than B/ℓ is $O(B)$: the remaining clients can contribute an additional cost of at most B, since at most ℓ such clients count towards the objective value of our solution. Thus, instead of bounding the cost of every set of ℓ clients, *our LP seeks to minimize the total assignment cost of clients having assignment cost larger than B/ℓ.*

More precisely, given a "guess" B of the optimal value, we consider the following LP. For $d \geq 0$, define $f_B(d) = d$ if $d > B/\ell$, and 0 otherwise. Throughout, i and j index points of $\mathcal{D}$.

$$\min \quad \sum_j \sum_i f_B(c_{ij}) x_{ij} \tag{P$_B$}$$

$$\text{s.t.} \quad \sum_i x_{ij} \geq 1 \qquad \forall j \tag{1}$$

$$0 \leq x_{ij} \leq y_i \qquad \forall i, j \tag{2}$$

$$\sum_i y_i \leq k. \tag{3}$$

Variable y_i indicates if facility i is open (i.e., i is chosen as a center), and x_{ij} indicates if client j is assigned to facility i. The first two constraints say that each client must be assigned to an open facility, and the third constraint encodes that at most k centers may be chosen.

An atypical aspect of our relaxation is that, while an integer solution corresponds to a solution to our problem, its objective value under (P$_B$) may *underestimate* the actual objective value; however, as alluded to above, the objective value of (P$_B$) is within an additive B of the actual objective value. Let OPT_B denote the optimal value of (P$_B$), and opt denote the optimal value of the ℓ-centrum problem.

Claim 2.1. *If $B \geq opt$, then $OPT_B \leq opt \leq B$.*

Proof. Let $(\tilde{x}, \tilde{y})$ be the integer point corresponding to an optimal solution. It is clear that $(\tilde{x}, \tilde{y})$ is feasible to (P$_B$). Also, there can be at most ℓ assignment costs that are larger than opt/ℓ, and hence at most ℓ assignment costs are larger than B/ℓ. Therefore, the objective value of $(\tilde{x}, \tilde{y})$ is at most opt. □

Claim 2.2. *Let $\vec{c}$ be an assignment-cost vector (where $\vec{c}_j$ is the assignment cost of j). Then, $cost(\ell; \vec{c}) \leq \sum_j f_B(\vec{c}_j) + B$.*

Proof. For any client j for which $\vec{c}_j$ is counted towards $cost(\ell; \vec{c})$ but $f_B(\vec{c}_j) = 0$, we have $\vec{c}_j \leq B/\ell$; there can be at most ℓ such clients, so the statement follows. □

The following claim shows that the weighted distances $\{f_B(c_{ij})\}$ satisfy an approximate form of triangle inequality.

Claim 2.3. *For any $B \geq 0$, we have: (i) $f_B(x) \leq f_B(y)$ if $x \leq y$; (ii) $\max\{f_B(x), f_B(y), f_B(z)\} \geq f_B\left(\frac{x+y+z}{3}\right)$ for any $x, y, z \geq 0$; and (iii) $3f_B(x/3) = f_{3B}(x)$ for any $x \geq 0$.*

Using binary search, we can find, within a $(1+\epsilon)$-factor, the smallest B such that $OPT_B \leq B$. Let $\overline{B}$ be this B. (Alternatively, we may enumerate all possible choices for opt in powers of $(1+\epsilon)$, and return the best solution among the solutions found for each B.) By Claim 2.1, we have that $\overline{B} \leq (1+\epsilon)opt$.

While (P$_B$) closely resembles the LP-relaxation for k-median, notice that the assignment costs $\{f_B(c_{ij})\}$ used in the objective of (P$_B$) *do not form a metric*. Despite this complication, we show that (P$_{\overline{B}}$) can be leveraged to obtain a solution of (actual, $cost(\ell; .)$-) cost $O(\overline{B})$. We describe two ways of obtaining such a guarantee, both of which are obtained via simple procedures and a clean analysis. The first is a primal-dual based algorithm based on the Jain-Vazirani (JV) template [6]. We Lagrangify (3) and move to the facility-location version where we may choose any number of centers but incur a fixed cost of (say) λ for each center we choose. By fine-tuning λ, we can find two solutions, one opening less than k centers, and the other opening more than k centers; rounding a convex combination of these solutions yields the final solution. This yields a 12-approximation algorithm. The second algorithm is based on LP-rounding, and shows that any α-approximation algorithm for k-median whose guarantee is with respect to the natural LP for k-median, can be used to obtain a solution of cost at most $2(\alpha+1)\overline{B}$.

Theorem 2.4. *We can obtain a solution to the ℓ-centrum problem of cost at most $\bigl(12+O(\epsilon)\bigr)\cdot\overline{B} \le \bigl(12+O(\epsilon)\bigr)\,opt$.*

Theorem 2.5. *Let (kmed-P) denote the k-median LP: $\min\bigl\{\sum_{j,i} c_{ij}x_{ij} \;:\; \text{(1)–(3)}\bigr\}$. Let $\mathcal{A}$ be an α-approximation algorithm for k-median whose approximation guarantee is proved relative to (kmed-P). We can obtain a solution to the ℓ-centrum problem of cost at most $2(\alpha+1)\overline{B}$. Thus, taking $\mathcal{A}$ to be the 3.25-approximation algorithm in [5], we obtain an $(8.5+\epsilon)$-approximation algorithm for the ℓ-centrum problem.*

In our algorithms and analysis, we have chosen to keep the exposition simple and not sought to overly optimize the constants. Although Theorem 2.4 yields a worse approximation guarantee, the underlying primal-dual algorithm and analysis are quite versatile and extend fairly easily to the setting with general weights. The remainder of this section is devoted to proving Theorem 2.4. We defer the proof of Theorem 2.5 to Appendix A.

2.1 Proof of Theorem 2.4

As noted earlier, the proof relies on the primal-dual method. The dual of ($\mathrm{P}_{\overline{B}}$) is as follows.

$$\max \quad \sum_j \alpha_j - k\cdot\lambda \tag{$\mathrm{D}_{\overline{B}}$}$$

$$\text{s.t.} \qquad \alpha_j \le f_{\overline{B}}(c_{ij}) + \beta_{ij} \qquad \forall i,j \tag{4}$$

$$\sum_j \beta_{ij} \le \lambda \qquad \forall i \tag{5}$$

$$\alpha,\lambda \ge 0.$$

Let $OPT := OPT_{\overline{B}}$ denote the optimal value of ($\mathrm{P}_{\overline{B}}$). We first fix λ and construct a solution that may open more than k centers but will have some near-optimality properties (see Theorem 2.6) as follows.

P1. **Dual-ascent.** Initialize $\mathcal{D}'=\mathcal{D}$, $\alpha_j=\beta_{ij}=0$ for all $i,j\in\mathcal{D}$, $F=\emptyset$. The clients in $\mathcal{D}'$ are called *active clients*. If $\alpha_j \ge f_{\overline{B}}(c_{ij})$, we say that j reaches i. (So if $c_{ij}\le \overline{B}/\ell$, then j reaches i from the very beginning.)

We repeat the following until all clients become inactive. Uniformly raise the α_js of all active clients, and the β_{ij}s for (i,j) such that $i\notin F$, j is active, and can reach i until one of the following events happen.

- Some client $j\in\mathcal{D}$ reaches some i (and previously could not reach i): if $i\in F$, we *freeze* j, and remove j from $\mathcal{D}'$.
- Constraint (5) becomes tight for some $i\notin F$: we add i to F; for every $j\in\mathcal{D}'$ that can reach i, we freeze j and remove j from $\mathcal{D}'$.

P2. **Pruning.** Pick a maximal subset T of F with the following property: for every $j\in\mathcal{D}$, there is at most one $i\in T$ such that $\beta_{ij}>0$.

P3. Return T as the set of centers, and assign every j to the nearest point in T, which we denote by $i(j)$.

Let $S=\{j : \exists i\in T \text{ s.t. } \beta_{ij}>0\}$.

Theorem 2.6. *The solution computed above satisfies $3\lambda|T| + \sum_{j\in S} f_{\overline{B}}(c_{i(j)j}) + \sum_{j\notin S} f_{3\overline{B}}(c_{i(j)j}) \le 3\sum_j \alpha_j$.*

Proof. The proof resembles the analysis of the JV primal-dual algorithm for facility location, but the subtlety is that we need to deal with the complication that the $\{f_{\overline{B}}(c_{ij})\}_{i,j\in\mathcal{D}}$ "distances" do not form a metric.

Observe that for every $i \in T$, every client $j \in S$ for which $\beta_{ij} > 0$ satisfies $i(j) = i$. So we have

$$\sum_{j \in S} 3\alpha_j \geq \sum_{j \in S}\Big(3\beta_{i(j)j} + f_{\overline{B}}(c_{i(j)j})\Big) = 3\lambda|T| + \sum_{j \in S} f_{\overline{B}}(c_{i(j)j}).$$

We show that for each client $j \notin S$, there is some $i'' \in T$ such that $f_{3\overline{B}}(c_{i''j}) \leq 3\alpha_j$, which will complete the proof. Let $i \in F$ be the facility that caused j to freeze, so $f_{\overline{B}}(c_{ij}) \leq \alpha_j$. If $i \in T$, then we are done. Otherwise, since T is maximal, there is some $i' \in T$ and some client $k \in S$ such that $\beta_{i'k}, \beta_{ik} > 0$. Notice that $\alpha_j \geq \alpha_k$, since α_j grows at least until the time point when i joins F, and α_k grows until at most this time point. Therefore, $f_{\overline{B}}(c_{ik}), f_{\overline{B}}(c_{i'k}) \leq \alpha_k \leq \alpha_j$. So by Claim 2.3, we have $f_{3\overline{B}}(c_{i'j}) \leq f_{3\overline{B}}(c_{i'k} + c_{ik} + c_{ij}) \leq 3\alpha_j$. $\square$

At $\lambda = 0$, the above algorithm will open a center at every point in $\mathcal{D}$, so open more than k centers. Using standard arguments, by performing binary search on λ, we can achieve one of the following two outcomes.

(a) Obtain some λ such that the above algorithm returns a solution T with $|T| = k$: in this case, Theorem 2.6 implies that $\sum_j f_{3\overline{B}}(c_{i(j)j}) \leq 3OPT$, and Claim 2.2 then implies that the $cost(\ell;.)$-cost of our solution is at most $3OPT + 3\overline{B} \leq 6\overline{B}$.

(b) Obtain $\lambda_1 < \lambda_2$ with $\lambda_2 - \lambda_1 < \frac{\epsilon\overline{B}}{n}$ such that letting T_1 and T_2 be the solutions returned for λ_1 and λ_2, we have $k_1 := |T_1| > k > k_2 := |T_2|$. We describe below the procedure for extracting a low-cost feasible solution from T_1 and T_2, and analyze it, which will complete the proof of Theorem 2.4.

Extracting a feasible solution from T_1 and T_2 in outcome (b). Let $a, b \geq 0$ be such that $ak_1 + bk_2 = k$, $a + b = 1$. Thus, a convex combination of T_1 and T_2 yields a feasible fractional solution that is sometimes called a *bipoint solution*, and our task is to round this into a feasible solution. Let (α_1, β_1), (α_2, β_2) denote the dual solutions obtained for λ_1 and λ_2 respectively. Let $i_1(j)$ and $i_2(j)$ denote the centers to which j is assigned in T_1 and T_2 respectively. Let $d_{1,j} = f_{3\overline{B}}(c_{i_1(j)j})$ and $d_{2,j} = f_{3\overline{B}}(c_{i_2(j)j})$. Let $C_1 := \sum_j d_{1,j}$ and $C_2 := \sum_j d_{2,j}$. Then,

$$\begin{aligned} aC_1 + bC_2 &\leq 3a\Big(\sum_j \alpha_{1,j} - k_1\lambda_1\Big) + 3b\Big(\sum_j \alpha_{2,j} - k_2\lambda_2\Big) \\ &\leq 3a\Big(\sum_j \alpha_{1,j} - k\lambda_2\Big) + 3b\Big(\sum_j \alpha_{2,j} - k\lambda_2\Big) + 3ak_1(\lambda_2 - \lambda_1) \leq 3OPT + 3\epsilon\overline{B} \end{aligned}$$

where the last inequality follows since $(\alpha_1, \beta_1, \lambda_2)$, $(\alpha_2, \beta_2, \lambda_2)$ are feasible solutions to ($\mathrm{D}_{\overline{B}}$). If $b \geq 0.5$, then T_2 yields a feasible solution of $cost(\ell;.)$-cost at most $C_2 + 3\overline{B} \leq 6OPT + (3+\epsilon)\overline{B}$. So suppose $a \geq 0.5$. The procedure for rounding the bipoint solution is as follows.

B1. **Clustering.** We first match facilities in T_2 with a subset of facilities in T_1 as follows. Initialize $\mathcal{D}' \leftarrow \mathcal{D}$, $A \leftarrow \emptyset$, and $M \leftarrow \emptyset$. We repeatedly pick the client $j \in \mathcal{D}'$ with minimum $d_{1,j} + d_{2,j}$ value, and add j to A. We add the tuple $(i_1(j), i_2(j))$ to M, remove from $\mathcal{D}'$ all clients k (including j) such that $i_1(k) = i_1(j)$ or $i_2(k) = i_2(j)$, and set $\sigma(k) = j$ for all such clients. Let $M_1 = M$ denote the matching so far. Next, for each unmatched $i \in T_2$, we pick an arbitrary unmatched facility $i' \in T_1$, and add (i', i) to M. Let F be the set of T_1-facilities that are matched, and $S := \{j \in \mathcal{D} : i_1(j) \in F\}$. Note that $|F| = |M| = k_2$.

B2. **Opening facilities.** We will open either all facilities in F, or all facilities in T_2 (which are always matched). Additionally, we will open $k - k_2$ facilities from $T_1 \setminus F$. We formulate the following LP to determine how to do this. Variable θ indicates if we open the facilities in F, and variables z_i for every

$i \in T_1 \setminus F$ indicate if we open facility i.

$$\min \quad \sum_{j \in S}\big(\theta d_{1,j} + (1-\theta)d_{2,j}\big) + \sum_{k \notin S}\big(z_{i_1(k)}d_{1,k} + (1-z_{i_1(k)})(d_{2,k} + d_{1,\sigma(k)} + d_{2,\sigma(k)})\big) \tag{R-P}$$

$$\text{s.t.} \quad \sum_{i \in T_1 \setminus F} z_i \leq k - k_2, \qquad \theta \in [0,1], \quad z_i \in [0,1] \ \ \forall i \in T_1 \setminus F. \tag{6}$$

The above LP is integral, and opening the facilities specified by an integral optimal solution (as discussed above) yields a solution of $cost(\ell;.)$-cost at most $15\overline{B}$. In Remark 2.7, we show that a slight modification yields an improved $cost(\ell;.)$-cost of at most $12\overline{B}$.

Analysis. It suffices to show that (R-P) has a fractional solution of small objective value, and that any integral solution yields a feasible solution to our problem whose $cost(\ell;.)$-cost is comparable to the objective value of (R-P).

For the former, we argue that setting $\theta = a$, $z_i = a$ for all $i \in T_1 \setminus F$ yields a feasible solution of objective value at most $2(aC_1 + bC_2)$. We have $\sum_{i \in T_1 \setminus F} z_i = a(k_1 - k_2) = k - k_2$. Every $j \in S$ contributes $ad_{1,j} + bd_{2,j}$ to the objective value of (R-P), which is also its contribution to $aC_1 + bC_2$. Consider $k \notin S$ with $\sigma(k) = j$, so $d_{1,j} + d_{2,j} \leq d_{1,k} + d_{2,k}$. Its contribution to the objective value of (R-P) is $ad_{1,k} + b(d_{2,k} + d_{1,j} + d_{2,j}) \leq (a+b)d_{1,k} + 2bd_{2,k}$, which is at most twice its contribution to $aC_1 + bC_2$.

For the latter, suppose we have an integral solution $(\tilde{\theta}, \tilde{z})$ to (R-P). For every $k \in S$, the assignment cost is at most $\tilde{\theta}d_{1,k} + (1-\tilde{\theta})d_{2,k} + 3\overline{B}/\ell$. Now consider $k \notin S$ with $\sigma(k) = j$. If $\tilde{z}_{i_1(k)} = 1$, it's assignment cost is at most $d_{1,k} + 3\overline{B}\ell$. Otherwise, it's assignment cost is at most $c_{i_2(k)k} + c_{i_1(j)j} + c_{i_2(j)j} \leq d_{2,k} + d_{1,j} + d_{2,j} + 9\overline{B}/\ell$. Thus, the $cost(\ell;.)$-cost of our solution is at most the objective value of $(\tilde{\theta}, \tilde{z}) + 9\overline{B}$, which is at most $2(aC_1 + bC_2) + 9\overline{B} \leq 6OPT + (9 + 3\epsilon)\overline{B} \leq \big(15 + O(\epsilon)\big)\overline{B}$.

Remark 2.7 (Improvement to the guarantee stated in Theorem 2.4). The following slightly modified way of opening facilities given an integral optimal solution $(\tilde{\theta}, \tilde{z})$ to (R-P) yields a solution of $cost(\ell;.)$-cost at most $12\overline{B}$.

As before, we open the facilities in $T_1 \setminus F$ specified by the $\tilde{z}_i$ variables that are 1. If $\tilde{\theta} = 1$, we open all the T_1-facilities in $M \setminus M_1$, and if $\tilde{\theta} = 0$, we open all the T_2-facilities in $M \setminus M_1$. For some clients $j \in A$, we may now open a facility at j (instead of at $i_1(j)$ or $i_2(j)$). For every $j \in A$, if $\tilde{\theta}d_{1,j} + (1-\tilde{\theta})d_{2,j} = 0$, then we open a facility at j; otherwise, we proceed as before, and open a facility at $i_1(j)$ if $\tilde{\theta} = 1$ and at $i_2(j)$ if $\tilde{\theta} = 0$.

To bound the cost, we first show that every $k \in S$ has assignment cost at most $\tilde{\theta}d_{1,k} + (1-\tilde{\theta})d_{2,k} + 6\overline{B}/\ell$. If a facility is opened in $\{k, i_1(k), i_2(k)\}$, then this clearly holds. Otherwise, it must be that $k \notin A$ and a facility is opened at $j = \sigma(k)$; taking $i = i_1(k)$ if $\tilde{\theta} = 1$ and $i_2(k)$ if $\tilde{\theta} = 0$, we have that $c_{ik} \leq \tilde{\theta}d_{1,k} + (1-\tilde{\theta})d_{2,k} + 3\overline{B}/\ell$ and $c_{ij} \leq 3\overline{B}/\ell$.

Now consider $k \notin S$ with $\sigma(k) = j$. If $\tilde{z}_{i_1(k)} = 1$, it's assignment cost is at most $d_{1,k} + 3\overline{B}\ell$. Otherwise, a facility is opened in $\{j, i_1(j), i_2(j)\}$. If a facility is opened in $\{j, i_1(j)\}$, then k's assignment cost is at most $c_{i_2(k)k} + c_{i_2(j)j} \leq d_{2,k} + d_{1,j} + d_{2,j} + 6\overline{B}/\ell$. Otherwise, it must be that $\tilde{\theta} = 1$ and $d_{1,j} = c_{i_1(j)j} > 3\overline{B}/\ell$; in this case, k' assignment cost is at most $c_{i_2(k)k} + c_{i_2(j)j} + c_{i_1(j)j} \leq (d_{2,k} + 3\overline{B}/\ell) + (d_{2,j} + 3\overline{B}/\ell) + d_{1,j}$. Thus, the $cost(\ell;.)$-cost of our solution is at most the objective value of $(\tilde{\theta}, \tilde{z}) + 6\overline{B}$, which is at most $2(aC_1 + bC_2) + 6\overline{B} \leq 6OPT + (6 + 3\epsilon)\overline{B} \leq \big(12 + O(\epsilon)\big)\overline{B}$. This concludes the proof of Theorem 2.4.

3 The general weighted case

We now consider the general setting, where we have $n = |\mathcal{D}|$ non-increasing nonnegative weights $w_1 \geq w_2 \geq \ldots \geq w_n \geq 0$, and the goal is to open k centers from $\mathcal{D}$ and assign each client $j \in \mathcal{D}$ to a center

$i(j) \in F$, so as to minimize

$$cost\big(w; \vec{c} := \{c_{i(j)j}\}_{j\in\mathcal{D}}\big) := w^T\vec{c}^{\downarrow} = \sum_{j=1}^{n} w_j \vec{c}_j^{\downarrow}.$$

By combining the ideas in Section 2 with an enumeration procedure due to [2], we obtain the following result.

Theorem 3.1. *We can obtain an $\big(18 + O(\epsilon)\big)$-approximation algorithm for ordered k-median that runs in time* $\mathrm{poly}\big((\frac{n}{\epsilon})^{1/\epsilon}\big)$.

The key again is to define suitable proxy costs analogous to the $f_B(c_{ij})$s for the setting with general weights. By defining these appropriately, it will be easy to argue that the primal-dual algorithm and its analysis extend to the setting with general weights, since essentially the only property that we use about $\{f_B(c_{ij})\}$ costs in Section 2 is that they satisfy Claim 2.3. A direct extension of $f_B(.)$, based on estimating the optimal $cost(w; .)$-cost and defining suitable thresholds, does not yield an $O(1)$-approximation.[1] Instead, we utilize a clever enumeration idea due to Aouad and Segev [2].

In Section 3.1, we describe this enumeration procedure using our notation, and restate the main claims in [2] in a simplified form. Next, in Section 3, we discuss how to adapt the ideas in Section 2.4 to the k-median problem for the proxy costs (given by (8)) that we obtain from Section 3.1. At the end of this section, we combine this ingredients to prove Theorem 3.1.

3.1 Proxy costs and the enumeration idea of [2]

Throughout, let $\vec{o}^{\downarrow}$ denote the assignment-cost vector corresponding to an optimal solution, whose coordinates are sorted in non-increasing order. So the optimal cost opt is $\sum_{i=1}^{n} w_i\vec{o}_i^{\downarrow}$. By a standard argument, we can perturb w to eliminate very small weights w_i: for $i \in [n]$, set $\widetilde{w}_i = w_i$ if $w_i \geq \frac{\epsilon w_1}{n}$, and $\widetilde{w}_i = 0$ otherwise.

Claim 3.2. *For any vector $v \in \mathbb{R}^n_+$, we have $(1-\epsilon)cost(w; v) \leq cost(\widetilde{w}; v) \leq cost(w; v)$.*

Proof. Since $\widetilde{w}_i \leq w_i$ for all $i \in [n]$, the upper bound on $cost(\widetilde{w}; v)$ is immediate. We have

$$cost(\widetilde{w}; v) = \sum_{i=1}^{n} \widetilde{w}_i v_i^{\downarrow} = cost(w; v) - \sum_{i\in[n]: w_i<\epsilon w_1/n} w_i v_i^{\downarrow} \geq cost(w; v) - \tfrac{\epsilon w_1}{n} \cdot n v_1^{\downarrow}. \quad \square$$

In the sequel, we always work with the $\widetilde{w}$-weights. We guess an estimate M of $\vec{o}_1^{\downarrow}$, and group distances in the range $\big[\frac{\epsilon M}{n}, M\big]$ (roughly speaking) by powers of $(1+\epsilon)$. Let T be the largest integer such that $\frac{\epsilon M}{n}(1+\epsilon)^T \leq M$. For $r = 0, \ldots, T$, we define the distance interval $I_r := \big(\frac{\epsilon M}{n}(1+\epsilon)^{T-r}, \frac{\epsilon M}{n}(1+\epsilon)^{T-r+1}\big]$. Note that there are at most $1 + \log_{1+\epsilon}\big(\frac{n}{\epsilon}\big) = O\big(\frac{1}{\epsilon}\log\frac{n}{\epsilon}\big)$ intervals.

Finally, we guess a non-increasing vector $w_0^{\mathsf{est}} \geq w_1^{\mathsf{est}} \geq \ldots \geq w_T^{\mathsf{est}}$, where the w_r^{est}s are powers of $(1+\epsilon)$ in the range $[\frac{\epsilon\widetilde{w}_1}{n}, \widetilde{w}_1(1+\epsilon))$. As argued in [2], there are only $\exp\big(O(\frac{1}{\epsilon}\log(\frac{n}{\epsilon}))\big) = O\big((\frac{n}{\epsilon})^{1/\epsilon}\big)$ choices for $w^{\mathsf{est}} := (w_0^{\mathsf{est}}, \ldots, w_T^{\mathsf{est}})$. The intention is for w_r^{est} to represent (within a $(1+\epsilon)$-factor) the average $\widetilde{w}$-weight of the set $\{i \in [n] : \vec{o}_i^{\downarrow} \in I_r\}$. More precisely, we would like w_r^{est} to estimate the following quantity, for all $r \in \{0, \ldots, T\}$.

$$w_r^{\mathsf{avg}} := \begin{cases} \big(\sum_{i\in[n]:\vec{o}_i^{\downarrow}\in I_r} \widetilde{w}_i\big)/|\{i \in [n] : \vec{o}_i^{\downarrow} \in I_r\}| & \text{if } \{i \in [n] : \vec{o}_i^{\downarrow} \in I_r\} \neq \emptyset; \\ \min\{\widetilde{w}_i : \vec{o}_i^{\downarrow} \in \bigcup_{s<r} I_s\} & \text{if } \bigcup_{s<r} I_s \neq \emptyset; \\ \widetilde{w}_1 & \text{otherwise.} \end{cases} \tag{7}$$

[1] It does however lead to an $O(\log n)$-approximation.

The following claim will be useful.

Claim 3.3. *For any $r \in \{0, \dots, T\}$, we have $w_r^{\mathsf{avg}} \geq \max\{\widetilde{w}_i : \vec{o}_i^{\downarrow} \notin \bigcup_{s \leq r} I_s\}$.*

Proof. If w_r^{avg} is defined by cases 1 or 2 of (7), then the inequality follows since for every $i' \in \bigcup_{s \leq r} I_r$ and $i \notin \bigcup_{s \leq r} I_s$, we have $\widetilde{w}_{i'} \geq \widetilde{w}_i$ (since $\vec{o}_{i'}^{\downarrow} \geq \vec{o}_i^{\downarrow}$). If w_r^{avg} is defined by case 3 of (7), then $w_r^{\mathsf{avg}} = \widetilde{w}_1$, and again, the inequality holds. $\square$

Given M and the corresponding intervals $I_0, \dots, I_T$, and the vector w^{est}, we can now finally define our proxy costs as follows. For $d \geq 0$ and $\gamma \geq 1$, define

$$g_{M,w^{\mathsf{est}}}(\gamma; d) = \begin{cases} \widetilde{w}_1(1+\epsilon)d & \text{if } d/\gamma > \frac{\epsilon M}{n}(1+\epsilon)^{T+1}; \\ w_r^{\mathsf{est}} d & \text{if } d/\gamma \in I_r \text{ (where } r \in \{0, \dots, T\}) \\ 0 & \text{if } d/\gamma \leq \frac{\epsilon M}{n}. \end{cases} \tag{8}$$

The above definition is essentially the scaled surrogate function in [2]. We abbreviate $g_{M,w^{\mathsf{est}}}(1; d)$ to $g_{M,w^{\mathsf{est}}}(d)$. The following two key lemmas are analogous to Claims 2.1 and 2.2, and show that for the right choice of M and w^{est}, evaluating the above proxy costs on an assignment-cost vector $\vec{c}$ yields a good estimate of the actual $cost(\widetilde{w}; .)$-cost of $\vec{c}$. Similar statements, albeit stated somewhat differently, are proved in [2].

Lemma 3.4 (adapted from [2])**.** *Suppose $M \geq \vec{o}_1^{\downarrow}$ and the w^{est} satisfies $w_r^{\mathsf{est}} \leq (1+\epsilon)w_r^{\mathsf{avg}}$ for all $r \in \{0, \dots, T\}$. Then, $\sum_{i=1}^n g_{M,w^{\mathsf{est}}}(\vec{o}_i^{\downarrow}) \leq (1+\epsilon)^2 cost(\widetilde{w}; \vec{o}^{\downarrow})$.*

Proof. Since $M \geq \vec{o}_1^{\downarrow}$, there is no i such that $\vec{o}_i^{\downarrow} > \frac{\epsilon M}{n}(1+\epsilon)^{T+1}$. Fix $r \in \{0, \dots, T\}$, and consider all $i \in [n]$ such that $\vec{o}_i^{\downarrow} \in I_r$. We have

$$\begin{aligned} \sum_{i \in [n]: \vec{o}_i^{\downarrow} \in I_r} g_{M,w^{\mathsf{est}}}(\vec{o}_i^{\downarrow}) = w_r^{\mathsf{est}} \sum_{i \in [n]: \vec{o}_i^{\downarrow} \in I_r} \vec{o}_i^{\downarrow} &\leq \frac{\epsilon M}{n}(1+\epsilon)^{T-r+1} \cdot w_r^{\mathsf{est}} \cdot \left|\{i \in [n] : \vec{o}_i^{\downarrow} \in I_r\}\right| \\ &\leq (1+\epsilon) \cdot \frac{\epsilon M}{n}(1+\epsilon)^{T-r+1} \cdot w_r^{\mathsf{avg}} \cdot \left|\{i \in [n] : \vec{o}_i^{\downarrow} \in I_r\}\right| \\ &= (1+\epsilon) \cdot \frac{\epsilon M}{n}(1+\epsilon)^{T-r+1} \cdot \sum_{i \in [n]: \vec{o}_i^{\downarrow} \in I_r} \widetilde{w}_i \leq (1+\epsilon)^2 \sum_{i \in [n]: \vec{o}_i^{\downarrow} \in I_r} \widetilde{w}_i \vec{o}_i^{\downarrow}. \end{aligned}$$

It follows that $\sum_{i=1}^n g_{M,w^{\mathsf{est}}}(\vec{o}_i^{\downarrow}) \leq (1+\epsilon)^2 cost(\widetilde{w}; \vec{o}^{\downarrow})$. $\square$

Lemma 3.5 (adapted from [2])**.** *Let $\gamma \geq 1$. Let $M \geq 0$, and suppose w^{est} satisfies $w_r^{\mathsf{avg}} \leq w_r^{\mathsf{est}}$ for all $r \in \{0, \dots, T\}$. Let $\vec{c}$ be an assignment-cost vector. Then, we have the upper bound $cost(\widetilde{w}; \vec{c}) \leq \sum_{i=1}^n g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i) + \gamma(1+\epsilon)cost(\widetilde{w}; \vec{o}^{\downarrow}) + \gamma\epsilon\widetilde{w}_1 M$.*

Proof. We have

$$cost(\widetilde{w}; \vec{c}) = \sum_{i=1}^n \widetilde{w}_i \vec{c}_i^{\downarrow} \leq \sum_{i=1}^n g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i) + \sum_{i:\ \widetilde{w}_i \vec{c}_i^{\downarrow} > g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow})} \widetilde{w}_i \vec{c}_i^{\downarrow}.$$

Consider some $i \in [n]$ for which $\widetilde{w}_i \vec{c}_i^{\downarrow} > g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow})$. It must be that $\vec{c}_i^{\downarrow}/\gamma \leq \frac{\epsilon M}{n}(1+\epsilon)^{T+1}$ as otherwise (see (8)), we have $g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow}) = (1+\epsilon)\widetilde{w}_1 \vec{c}_i^{\downarrow} > \widetilde{w}_i \vec{c}_i^{\downarrow}$. If $g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow}) = 0$, then we have $\widetilde{w}_i \vec{c}_i^{\downarrow}/\gamma \leq \widetilde{w}_i \cdot \frac{\epsilon M}{n} \leq \widetilde{w}_1 \cdot \frac{\epsilon M}{n}$.

Otherwise, we claim that $\vec{c}_i^{\downarrow}/\gamma \leq (1+\epsilon)\vec{o}_i^{\downarrow}$. Suppose not. Suppose $\vec{c}_i^{\downarrow}/\gamma \in I_r$, where $r \in \{0, \ldots, T\}$. Since $\frac{\vec{c}_i^{\downarrow}/\gamma}{\vec{o}_i^{\downarrow}} > (1+\epsilon)$, we have that $\vec{o}_i^{\downarrow} \notin \bigcup_{s \leq r} I_s$. So by Claim 3.3, we have $w_r^{\mathsf{avg}} \geq \widetilde{w}_i$. Hence, $g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow}) = w_r^{\mathsf{est}} \vec{c}_i^{\downarrow} \geq w_r^{\mathsf{avg}} \vec{c}_i^{\downarrow} \geq \widetilde{w}_i \vec{c}_i^{\downarrow}$, which contradicts our assumption that $\widetilde{w}_i \vec{c}_i^{\downarrow} > g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow})$.

Putting everything together, we have that $\sum_{i: \widetilde{w}_i \vec{c}_i^{\downarrow} > g_{M,w^{\mathsf{est}}}(\gamma; \vec{c}_i^{\downarrow})} \widetilde{w}_i \vec{c}_i^{\downarrow} \leq n\gamma \widetilde{w}_1 \cdot \frac{\epsilon M}{n} + \gamma(1+\epsilon) \sum_{i \in [n]} \widetilde{w}_i \vec{o}_i^{\downarrow}$, which proves the lemma. □

Finally, we show that $g_{M,w^{\mathsf{est}}}$ satisfies the analogue of Claim 2.3, which will be crucial in arguing that our algorithms and analysis from Section 2.4 carry over and allow us to solve, in an approximate sense, the k-median problem with the $\{g_{M,w^{\mathsf{est}}}(c_{ij})\}$ proxy costs.

Lemma 3.6. *For any $\gamma \geq 1$, $M \geq 0$, and w^{est}, we have: (i) $g_{M,w^{\mathsf{est}}}(\gamma; x) \leq g_{M,w^{\mathsf{est}}}(\gamma; y)$ if $x \leq y$; and (ii) $3 \max\{g_{M,w^{\mathsf{est}}}(\gamma; x), g_{M,w^{\mathsf{est}}}(\gamma; y), g_{M,w^{\mathsf{est}}}(\gamma; z)\} \geq g_{M,w^{\mathsf{est}}}(3\gamma; x+y+z)$ for any $x, y, z \geq 0$.*

Proof. Part (i) follows readily from the definition (8). Part (ii) follows from part (i) by noting that $g_{M,w^{\mathsf{est}}}(3\gamma; x+y+z) = 3g_{M,w^{\mathsf{est}}}\big(\gamma; \frac{x+y+z}{3}\big)$. □

3.2 Solving the k-median problem with the $\{g_{M,w^{\mathsf{est}}}(c_{ij})\}$ proxy costs

We now work with a fixed guess M, w^{est}, and give an algorithm for finding a near-optimal k-median solution with the $\{g_{M,w^{\mathsf{est}}}(c_{ij})\}$ proxy costs. Our algorithm and analysis will be quite similar to the one in Section 2.4. The primal and dual LPs we consider are the same as (P_B) and ($\mathrm{D}_{\overline{B}}$), except that we replace all occurrences of $f_B(c_{ij})$ and $f_{\overline{B}}(c_{ij})$ with $g_{M,w^{\mathsf{est}}}(c_{ij})$. Let $OPT_{M,w^{\mathsf{est}}}$ denote the optimal value of this LP.

The primal-dual algorithm for a given center-cost λ (steps P1–P3 in Section 2.4) is unchanged. The analysis also is essentially identical, since, previously, we only relied on the fact that the proxy costs satisfy an approximate triangle inequality, which is also true here (Lemma 3.6). We state the guarantee from the primal-dual algorithm slightly differently, in the form suggested by part (ii) of Lemma 3.6. The proof of the following theorem simply mimics the proof of Theorem 2.6.

Theorem 3.7. *For any $\lambda \geq 0$, the primal-dual algorithm (P1)–(P3) returns a set T of centers, an assignment $i(j) \in T$ for every $j \in \mathcal{D}$, and a dual feasible solution (α, β, λ) such that $3\lambda|T| + \sum_j g_{M,w^{\mathsf{est}}}(3; c_{i(j)j}) \leq 3\sum_j \alpha_j$.*

Given Theorem 3.7, we can use binary search on λ, to either obtain: (a) some λ such for which we return a solution T with $|T| = k$; or (b) $\lambda_1 < \lambda_2$ with $\lambda_2 - \lambda_1 < \frac{\epsilon \widetilde{w}_1 M}{n}$ such that letting T_1 and T_2 be the solutions returned for λ_1 and λ_2, we have $k_1 := |T_1| > k > k_2 := |T_2|$. In case (a), Theorem 3.7 implies that $\sum_j g_{M,w^{\mathsf{est}}}(3; c_{i(j)j}) \leq 3OPT_{M,w^{\mathsf{est}}}$. In case (b), we again extract a low-cost feasible solution from T_1 and T_2 by rounding the bipoint solution given by their convex combination. As before, $a, b \geq 0$ be such that $ak_1 + bk_2 = k$, $a + b = 1$. Let (α_1, β_1), (α_2, β_2) denote the dual solutions obtained for λ_1 and λ_2 respectively. Let $i_1(j)$ and $i_2(j)$ denote the centers to which j is assigned in T_1 and T_2 respectively. Let $d_{1,j} = g_{M,w^{\mathsf{est}}}(3; c_{i_1(j)j})$ and $d_{2,j} = g_{M,w^{\mathsf{est}}}(3; c_{i_2(j)j})$. Let $C_1 := \sum_j d_{1,j}$ and $C_2 := \sum_j d_{2,j}$. Similar to before, we have $aC_1 + bC_2 \leq 3OPT_{M,w^{\mathsf{est}}} + 3\epsilon \widetilde{w}_1 M$. The procedure for rounding this bipoint solution requires only minor changes to steps B1, B2 in Section 2.4, as we now describe.

Rounding the bipoint solution obtained from T_1, T_2. If $b \geq 1/3$, then T_2 yields a feasible solution with $\sum_j g_{M,w^{\mathsf{est}}}(3; c_{i_2(j)j}) = C_2 \leq 9OPT_{M,w^{\mathsf{est}}} + 9\epsilon \widetilde{w}_1 M$. So suppose $a \geq 2/3$.

G1. **Clustering.** We match facilities in T_2 with a subset of facilities in T_1 as follows. Initialize $\mathcal{D}' \leftarrow \mathcal{D}$, $A \leftarrow \emptyset$, and $M \leftarrow \emptyset$. We repeatedly pick the client $j \in \mathcal{D}'$ with minimum $\max\{d_{1,j}, d_{2,j}\}$ value, and add j to A. (***This is the only change, compared to step B1.***) We add the tuple $(i_1(j), i_2(j))$ to M, remove from $\mathcal{D}'$ all clients k (including j) such that $i_1(k) = i_1(j)$ or $i_2(k) = i_2(j)$, and set $\sigma(k) = j$ for all such clients. Let $M_1 = M$ denote the matching so far. Next, for each unmatched $i \in T_2$, we pick an arbitrary unmatched facility $i' \in T_1$, and add (i', i) to M. Let F be the set of T_1-facilities that are matched, and $S := \{j \in \mathcal{D} : i_1(j) \in F\}$. Note that $|F| = |M| = k_2$.

G2. **Opening facilities.** This is almost identical to step B2, except that we decide which facilities to open by now solving the following LP.

$$\begin{aligned} \min \quad & \sum_{j \in S} \bigl(\theta d_{1,j} + (1-\theta) d_{2,j}\bigr) + \sum_{k \notin S} \bigl(z_{i_1(k)} d_{1,k} + (1 - z_{i_1(k)}) \cdot 3\max\{d_{1,k}, d_{2,k}\}\bigr) \qquad \text{(GR-P)} \\ \text{s.t.} \quad & \sum_{i \in T_1 \setminus F} z_i \le k - k_2, \qquad \theta \in [0,1], \quad z_i \in [0,1] \;\; \forall i \in T_1 \setminus F. \end{aligned}$$

Let $(\tilde{\theta}, \tilde{z})$ be an optimal integral solution to (GR-P). As before, if $\tilde{\theta} = 1$, we open all facilities in F, and otherwise, all facilities in T_2. We also the facilities from $T_1 \setminus F$ for which $\tilde{z}_i = 1$.

To analyze this, we first show that setting $\theta = a$, $z_i = a$ for all $i \in T_1 \setminus F$ yields a feasible solution to (GR-P) of objective value at most $3(aC_1 + bC_2)$. We have $\sum_{i \in T_1 \setminus F} z_i = a(k_1 - k_2) = k - k_2$. Every $j \in S$ contributes $ad_{1,j} + bd_{2,j}$ to the objective value of (GR-P). Consider $k \notin S$. Its contribution to the objective value of (GR-P) is

$$ad_{1,k} + 3b\max\{d_{1,k}, d_{2,k}\} = \max\{(a+3b)d_{1,k}, ad_{1,k} + 3bd_{2,k}\} \le 3(ad_{1,k} + bd_{2,k})$$

where the inequality follows since $a + 3b \le 3a$ when $a \ge 2/3$. Thus, for every $j \in \mathcal{D}$, its contribution to the objective value of (GR-P) is at most thrice its contribution to $aC_1 + bC_2$.

Suppose $\vec{c}$ is the assignment-cost vector resulting from $(\tilde{\theta}, \tilde{z})$. We show that $\sum_j g_{M,w^{\mathsf{est}}}(9; \vec{c}_j)$ is at most the objective value of $(\tilde{\theta}, \tilde{z})$ under (GR-P). For every $k \in S$, we have $g_{M,w^{\mathsf{est}}}(9; \vec{c}_k) \le g_{M,w^{\mathsf{est}}}(3; \vec{c}_k) \le \tilde{\theta} d_{1,k} + (1 - \tilde{\theta}) d_{2,k}$. Now consider $k \notin S$ with $\sigma(k) = j$, so $\max\{d_{1,j}, d_{2,j}\} \le \max\{d_{1,k}, d_{2,k}\}$. If $\tilde{z}_{i_1(k)} = 1$, then $g_{M,w^{\mathsf{est}}}(9; \vec{c}_k) \le g_{M,w^{\mathsf{est}}}(3; \vec{c}_k) \le d_{1,k}$. Otherwise, $\vec{c}_k \le c_{i_2(k)k} + c_{i_1(j)j} + c_{i_2(j)j}$, and so by Lemma 3.6, we have

$$\begin{aligned} g_{M,w^{\mathsf{est}}}(9; \vec{c}_k) &\le g_{M,w^{\mathsf{est}}}(9; c_{i_2(k)k} + c_{i_1(j)j} + c_{i_2(j)j}) \\ &\le 3\max\{g_{M,w^{\mathsf{est}}}(3; c_{i_2(k)}), g_{M,w^{\mathsf{est}}}(3; c_{i_1(j)j}), g_{M,w^{\mathsf{est}}}(3; c_{i_2(j)j})\} \le 3\max\{d_{1,k}, d_{2,k}\}. \end{aligned}$$

So in every case, $g_{M,w^{\mathsf{est}}}(9; \vec{c}_k)$ is bounded by the contribution of k to the objective value of $(\tilde{\theta}, \tilde{z})$. Thus, we have proved the following theorem.

Theorem 3.8. *For any $M \ge 0$, w^{est}, we can obtain a solution opening k centers whose assignment-cost vector $\vec{c}$ satisfies $\sum_j g_{M,w^{\mathsf{est}}}(9; \vec{c}_j) \le 9\, OPT_{M,w^{\mathsf{est}}} + 9\epsilon \widetilde{w}_1 M$.*

Proof of Theorem 3.1. The proof follows by combining Theorem 3.8, Lemmas 3.4 and 3.5, and Claim 3.2.

Recall that $\vec{o}^{\downarrow}$ is the assignment-cost vector corresponding to an optimal solution with coordinates sorted in non-increasing order, and $opt = \sum_{i=1}^{n} w_i \vec{o}_i^{\downarrow}$ is the optimal cost.

There are only n^2 choices for M, and $O\bigl((\frac{n}{\epsilon})^{1/\epsilon}\bigr)$ choices for w^{est}, so we may assume that in polynomial time, we have obtained $M = \vec{o}_1^{\downarrow}$, and w_r^{est}s satisfying $w_r^{\mathsf{avg}} \le w_r^{\mathsf{est}} \le (1+\epsilon) w_r^{\mathsf{avg}}$ for all $r \in \{0, \ldots, T\}$. By Lemma 3.4, we know that $OPT_{M,w^{\mathsf{est}}} \le (1+\epsilon)^2 cost(\widetilde{w}; \vec{o}^{\downarrow}) \le (1+\epsilon)^2 opt$. Let $\vec{c}$ be the assignment-cost vector of the solution returned by Theorem 3.8 for this M, w^{est}. Combining Theorem 3.8, Lemma 3.5, and Claim 3.2, we obtain that

$$\begin{aligned} (1-\epsilon) cost(w; \vec{c}) \le cost(\widetilde{w}; \vec{c}) &\le \bigl(9\, OPT_{M,w^{\mathsf{est}}} + 9\epsilon \widetilde{w}_1 M\bigr) + 9(1+\epsilon) cost(\widetilde{w}; \vec{o}^{\downarrow}) + 9\epsilon \widetilde{w}_1 M \\ &\le 9(1+\epsilon)^2 opt + 9\, opt + O(\epsilon)\, opt = \bigl(18 + O(\epsilon)\bigr) opt. \end{aligned}$$

References

[1] S. Alamdari and D. Shmoys. A bicriteria approximation algorithm for the k-center and k-median problems To appear in *Proceedings of WAOA*, 2017

[2] A. Aouad and D. Segev. The ordered k-median problem: surrogate models and approximation algorithms. *In submission.*

[3] J. Byrka, K. Sornat, and J.Spoerhase. Constant-factor approximation for ordered k-median. *CS arXiv, arXiv:1711.01972 [cs.DS]*, Nov 6, 2017.

[4] M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys. A constant-factor approximation algorithm for the k-median problem. *Journal of Computer and System Sciences*, 65(1):129–149, 2002.

[5] M. Charikar and S. Li. A dependent LP-rounding approach for the k-median problem. In *Proceedings of the 39th ICALP*, pages 194–205, 2012.

[6] K. Jain and V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and lagrangian relaxation. *Journal of the ACM*, 48(2):274–296, 2001.

[7] G. Laporte, S. Nickel, and F. S. da Gama. *Location Science*, Springer, 2015.

[8] S. Nickel and J. Puerto. *Location Theory: A Unified Approach*, Springer Science & Business Media, 2005.

[9] J. Puerto and A. Tamir. Locating tree-shaped facilities using the ordered median objective. *Mathematical Programming*, 102(2):313–338, 2005.

[10] A. Tamir. The k-centrum multi-facility location problem. *Discrete Applied Mathematics*, 109(3):293–307, 2001.

A Proof of Theorem 2.5

Recall that $\overline{B} \leq (1+\epsilon)opt$ is such that $OPT_{\overline{B}} \leq \overline{B}$, and $\mathcal{A}$ is an α-approximation algorithm for k-median whose approximation guarantee is proved relative to the natural LP ($k$med-P) for k-median. Let (x, y) denote an optimal solution to ($\mathrm{P}_{\overline{B}}$), whose objective value is $OPT := OPT_{\overline{B}}$. Define $\mathsf{LP}_j := \sum_i f_{\overline{B}}(c_{ij})x_{ij}$ to be the cost incurred for client j by the LP ($\mathrm{P}_{\overline{B}}$).

Our rounding algorithm is quite simple: we perform clustering and demand consolidation to merge clients that are (roughly speaking) within distance $\overline{B}/\ell$ of each other. This *reduces* our instance to a k-median instance, and we then run algorithm $\mathcal{A}$ on this instance.

R1. **Clustering and demand consolidation.** Set $d'_j \leftarrow 0$ for every j. Consider the clients in increasing order of LP_j. For each client k encountered, if there exists a client j such that $d'_j > 0$ and $c_{jk} \leq 2\overline{B}/\ell$, set $d'_j \leftarrow d'_j + 1$, otherwise set $d'_k \leftarrow 1$. Let $D' = \{j \in \mathcal{D} : d'_j > 0\}$. Each client in D' is a cluster center. For $k \in \mathcal{D} \setminus D'$, we set $\sigma(k) = j$, if k's demand was moved to j above; we set $\sigma(j) = j$ for all $j \in D'$.

R2. **Running k-median.** Consider the k-median instance $\mathcal{I}'$ consisting of the weighted point-set $\{d'_j\}_{j \in D'}$ (and the c_{ij}-distances between these points). Note that the points in $\mathcal{D} \setminus D'$ do not appear in $\mathcal{I}'$. We run algorithm $\mathcal{A}$ to solve instance $\mathcal{I}'$ and obtain our k centers.

Analysis. Let $OPT' := \sum_{j \in D', i} d'_j f_{\overline{B}}(c_{ij}) x_{ij}$ denote the LP-cost of (x, y) for the modified instance consisting of the cluster centers. For each $j \in D'$, define F_j to be all the points $i \in \mathcal{D}'$ such that j is the point in D' closest to i, that is, $F_j := \{i \in \mathcal{D}' : c_{ij} = \min_{j' \in D'} c_{ij'}\}$. We break ties arbitrarily, so the F_js are disjoint.

Claim A.1. *(i) If $j, k \in D'$, then $c_{jk} > 2\overline{B}/\ell$, and (ii) $OPT' \leq OPT$.*

Proof. Suppose k was considered after j. Then $d'_j > 0$ at this time, as d'_j becomes positive when j is added to D'. So if $c_{jk} \leq 2\overline{B}/\ell$ then d'_k would remain at 0, giving a contradiction. It is clear that if we move the demand of client k to client j, then $\mathsf{LP}_j \leq \mathsf{LP}_k$ and $c_{jk} \leq 2\overline{B}/\ell$. So $OPT' = \sum_j d'_j \mathsf{LP}_j \leq \sum_j \mathsf{LP}_j = OPT$. □

Lemma A.2. *There is a fractional solution to (kmed-P) for the k-median instance $\mathcal{I}'$ of objective value at most $2OPT'$.*

Proof. Consider the following fractional solution. For each $j \in D'$, set $X_{jj} = \sum_{i \in F_j} y_i = Y_j$; for every distinct $j, j' \in D'$, set $X_{j'j} = \sum_{i \in F_{j'}} x_{ij}$. It is easy to verify that (X, Y) is a feasible solution to ($k$med-P). Since $y_i \geq x_{ij}$ for all $i, j \in \mathcal{D}'$, we have $\sum_{j' \in D'} X_{j'j} \geq \sum_{i \in \mathcal{D}'} x_{ij} \geq 1$ for every $j \in D'$, and $X_{j'j} \leq Y_{j'}$ for every distinct $j, j' \in \mathcal{D}'$; also $\sum_{j \in D'} Y_j = \sum_{i \in \mathcal{D}'} y_i \leq k$.

We now bound the objective value of (X, Y) in ($k$med-P) (for the weighted point set in $\mathcal{I}'$). Observe that for any $j \in D'$ and any $i \in F_{j'}$, where $j' \neq j$, we have $c_{ij} > \overline{B}/\ell$, as otherwise $c_{jj'} \leq 2\overline{B}/\ell$, contradicting part (i) of Claim A.1. Therefore, $f_{\overline{B}}(c_{ij}) = c_{ij}$, and $c_{jj'} \leq 2c_{ij} \leq 2f_{\overline{B}}(c_{ij})$. So we have

$$\sum_{j, j' \in D'} d'_j c_{jj'} X_{jj'} \leq \sum_{j, j' \in D' : j \neq j'} \sum_{i \in F_{j'}} 2d'_j f_{\overline{B}}(c_{ij}) x_{ij} \leq \sum_{j \in D'} \sum_i 2d'_j f_{\overline{B}}(c_{ij}) x_{ij} = 2OPT'. \quad \square$$

Lemma A.3. *Any (integer) solution to the k-median instance $\mathcal{I}'$ of cost C, yields a solution to the original ℓ-centrum instance of cost$(\ell; .)$-cost of at most $C + 2\overline{B}$.*

Proof. For $j \in D'$, let $i(j) \in D'$ denote the facility to which j is assigned in the solution to $\mathcal{I}'$. For any $k \in \mathcal{D} \setminus D'$ with $\sigma(k) = j$, its assignment cost for the original instance is at most $c_{i(j)k} \leq c_{i(j)j} + 2\overline{B}/\ell$. Thus, the assignment cost of any set of ℓ clients of the original instance is at most $C + 2\overline{B}$. □

Theorem 2.5 follows immediately from Lemmas A.2 and A.3 and part (ii) of Claim A.1: the $cost(\ell; .)$-cost of the solution obtained is at most $\alpha(2OPT') + 2\overline{B} \leq 2\alpha OPT + 2\overline{B} \leq 2(\alpha + 1)\overline{B}$.