

Constant-Factor Approximation for ORDERED k -MEDIAN

Jarosław Byrka*

Krzysztof Sornat†

Joachim Spoerhase‡

Abstract

We study the ORDERED k -MEDIAN problem, in which the solution is evaluated by first sorting the client connection costs and then multiplying them with a predefined non-increasing weight vector (higher connection costs are taken with larger weights). Since the 1990s, this problem has been studied extensively in the discrete optimization and operations research communities and has emerged as a framework unifying many fundamental clustering and location problems such as k -MEDIAN and k -CENTER. This generality, however, renders the problem intriguing from the algorithmic perspective and obtaining non-trivial approximation algorithms was an open problem even for simple topologies such as trees. Recently, Aouad and Segev were able to obtain an $\mathcal{O}(\log n)$ approximation algorithm for ORDERED k -MEDIAN using a sophisticated local-search approach and the concept of surrogate models thereby extending the result by Tamir (2001) for the case of a rectangular weight vector, also known as k -FACILITY p -CENTRUM. The existence of a constant-factor approximation algorithm, however, remained open even for the rectangular case.

In this paper, we provide an LP-rounding constant-factor approximation algorithm for the ORDERED k -MEDIAN problem.

We obtain this result by revealing an interesting connection to the classic k -MEDIAN problem. We first provide a new analysis of the rounding process by Charikar and Li (2012) for k -MEDIAN, when applied to a fractional solution obtained from solving an LP relaxation over a non-metric, truncated cost vector, resulting in an elegant 15-approximation for the rectangular case. In our analysis, the connection cost of a single client is partly charged to a deterministic budget related to a combinatorial bound based on guessing, and partly to a budget whose expected value is bounded with respect to the fractional LP-solution. This approach allows us to limit the problematic effect of the variance of individual client connection costs on the value of the ranking-based objective function of ORDERED k -MEDIAN. Next, we analyze objective-oblivious clustering, which allows us to handle multiple rectangles in the weight vector and obtain a constant-factor approximation for the case of $\mathcal{O}(1)$ rectangles. Then, we show that a simple weight bucketing can be applied to the general case resulting in $\mathcal{O}(\log n)$ rectangles and hence in a constant-factor approximation in quasi-polynomial time. Finally, with a more involved argument, we show that also the clever distance bucketing by Aouad and Segev can be combined with the objective-oblivious version of our LP-rounding for the rectangular case, and that it results in a true, polynomial time, constant-factor approximation algorithm.

*University of Wrocław, Wrocław, Poland, jby@cs.uni.wroc.pl.

†University of Wrocław, Wrocław, Poland, krzysztof.sornat@cs.uni.wroc.pl.

‡University of Wrocław, Wrocław, Poland, joachim.spoerhase@uni-wuerzburg.de.

1 Introduction

Clustering a given set of objects into k groups that display a certain internal proximity is a profound combinatorial optimization setting. In a typical setup, we represent the objects as points in a metric space and evaluate the quality of the clustering by a certain function of distances within clusters. If clusters have centers and the objective is to minimize the total distance from objects to their cluster centers, we call the resulting optimization problem k -MEDIAN. If the objective is to minimize the maximal distance to a cluster center, then we talk about the k -CENTER problem. These two approaches to clustering represent the two extremes in their dependence on the variance between the individual connection costs in the evaluated solution. Several intermediate approaches have been studied such as minimizing the sum of squared connection costs known as the k -MEANS problem.

In this paper, we study the ORDERED k -MEDIAN problem where the connection costs are sorted non-increasingly and a non-increasing weight vector is applied to flexibly penalize the desired fraction of the highest costs. There is a large body of literature on this problem because it naturally unifies many of the most fundamental clustering and location problems such as k -MEDIAN, k -CENTER, k -CENTDIAN, and k -FACILITY p -CENTRUM (see below for definitions). We refer to the book of Nickel and Puerto [31] dedicated to ORDERED k -MEDIAN problems for an extensive overview. See also below for a selection of related works and also applications in multi-objective optimization and robust optimization.

The generality of ORDERED k -MEDIAN renders it intriguing from the computational perspective [3]. For example, whereas k -MEDIAN and k -CENTER can be solved efficiently on trees by dynamic programming such approaches seem to fail for ORDERED k -MEDIAN due to the lack of separability properties [38]. Regarding approximability in general metric spaces, constant-factor approximation algorithms are long known for k -MEDIAN [13] and k -CENTER [21]. In contrast, not even non-trivial *super-constant* approximability results were known for ORDERED k -MEDIAN until very recently and even developing constant-factor approximations for seemingly simple topologies such as trees turned out non-trivial [3]. In particular, due to the non-linearity of the objective function there seems to be no obvious way to apply tools such as metric tree embeddings [3, 5]. To demonstrate the highly non-local and dependent structure of the objective function note that even if the clusters are given the selection of the cluster centers cannot be made solely on a per-cluster basis but depends on the decision in other clusters due to the ranking of distances in the objective.

Related works for ORDERED k -MEDIAN. The problem generalizes many fundamental clustering and location problems [13, 22, 4, 28, 11, 21, 42, 41] such as the above-mentioned k -MEDIAN, k -CENTER problems, the k -CENTDIAN problem where the objective is a convex combination of the k -MEDIAN and the k -CENTER objective, or the k -FACILITY p -CENTRUM problem where the objective accounts for the p highest connection costs. The ordered median objective function has also been considered in robust optimization [37, 7, 8, 6] and multi-objective optimization [19]. For a comprehensive overview we refer to the books [31, 27] on ORDERED k -MEDIAN problems and to dedicated works [30, 29, 10, 32].

Due to the above-outlined difficulties in obtaining algorithmic results for the general problem, structural properties of continuous network spaces have been studied [35, 16, 18, 39] where facilities may be placed at interior points on edges, single-facility models [30, 17, 24, 18], and multi-facility models on special topologies such as trees [23, 41, 37]. Also, integer programming formulations [29], branch-and-bound methods [9, 34], heuristics [15, 40, 33, 26], and related location models [36, 38] have been studied. For a survey on the topic we refer to [31].

Although the problem has received much attention in the discrete optimization and operations research literature, obtaining any non-trivial approximation algorithm for ORDERED k -MEDIAN

had been an intriguing open question until very recently, when Aouad and Segev [3] were able to devise an $\mathcal{O}(\log n)$ approximation algorithm for the problem using a sophisticated local-search approach and the concept of *surrogate models*.

Below, we list approximability results for certain specific objective functions that fall under the framework of ORDERED k -MEDIAN or that are closely related and where approximability has been studied for general metrics.

Approximation algorithms for k -MEDIAN, k -CENTER, and k -MEANS. k -MEDIAN admits constant-factor approximations via local-search [4], or a direct rounding of the standard LP [14]. The current best ratio of $(2.675 + \epsilon)$ [11] is obtained by combining a primal-dual algorithm [22], and a nontrivial rounding of a so-called bi-point solution based on a preprocessing introduced in [28].

The situation with the k -CENTER setting is simpler. A simple 2-approximation¹ is obtained via guessing the longest connection distance in the optimal solution [21], and this is tight assuming $P \neq NP$. Notably, by contrast to the k -MEDIAN setting, the most natural LP for k -CENTER has unbounded integrality gap.

Also k -MEANS admits a constant-factor approximation. The $(9 + \epsilon)$ -approximation algorithm for EUCLIDEAN k -MEANS [25] can be shown to provide 25-approximation in general metrics. The recent work of Ahmadian et al. [1] decreases these ratios to 6.375 and 9, respectively.

Approximation algorithms for further specific objective functions. A special case of ORDERED k -MEDIAN that we call RECTANGULAR ORDERED k -MEDIAN was considered by Tamir [41] (who called it k -FACILITY p -CENTRUM). In this setting, we have to open exactly k facilities and the objective function is just a sum of p largest client connection costs. He gives polynomial time algorithms that solve the problem (optimally) on path and tree graphs and using tree-embedding gives a $\mathcal{O}(\log n)$ -approximation for this case. Obtaining a constant-factor approximation for RECTANGULAR ORDERED k -MEDIAN, however, has been an open problem [41, 3].

One should also notice at least two further, very recent works combining the k -CENTER objective and the k -MEDIAN objective. First, Alamdari and Shmoys [2] considered a bicriteria approximation algorithm for the k -CENTER and k -MEDIAN problems, i.e., the objective function is a linear combination of two objectives that are maximal connection cost in use and sum of all used connection costs. This problem is known as k -CENTDIAN [42]. They obtained polynomial time bicriteria approximation of $(4, 8)$, where the first factor is in respect to the k -CENTER objective and the second factor is in respect to the k -MEDIAN objective. (Alamdari and Shmoys note, however, that the two problems k -MEDIAN and k -CENTER are *not* approximable simultaneously.) Also k -CENTDIAN is a special case of ORDERED k -MEDIAN. The second recent work combining the k -MEDIAN and the k -CENTER objective is the work of Haris et al. [20] who propose a method to select k facilities that deterministically guarantees each client to have a connection within a certain fixed radius but also provides a stronger per client bound on cost expectation.

Relation to the work of Chakrabarty and Swamy [12]. Soon after the submission of our paper, Chakrabarty and Swamy [12] announced constant-factor approximation algorithms for RECTANGULAR ORDERED k -MEDIAN and also for (general) ORDERED k -MEDIAN. The part of their argument for RECTANGULAR ORDERED k -MEDIAN appears to be obtained independently. Instead of the LP-rounding process of Charikar and Li [14], they either use a primal-dual approach or a black-box reduction to k -MEDIAN.

¹In the setting where cluster centers are selected from a different set the method of [21] gives a tight 3-approximation.

1.1 Our Results and Techniques

Our main result is an LP-rounding constant-factor approximation algorithm for the ORDERED k -MEDIAN problem.

We are not aware of a LP relaxation for ORDERED k -MEDIAN with bounded integrality gap. In our approach we guess a *reduced cost function* roughly mimicking the weighting of distances in an optimum solution and solve the natural LP relaxation for k -MEDIAN under this reduced cost function (rather than under the original metric). Subsequently, we round this solution via a dependent LP rounding process by Charikar and Li [14] for k -MEDIAN operating on the original (unweighted) metrics.

The challenge and our main technical contribution consists in analyzing the approximation performance of this approach. In the original analysis of Charikar and Li [14] for the k -MEDIAN objective, a per-client bound on the expected connection cost of this client with respect to its fractional connection cost is established. The global approximation ratio is then obtained by linearity of expectation. The above-described non-linear, ranking-based character of the objective of ORDERED k -MEDIAN poses an obstacle to apply an analogous reasoning also in our more general setting as the actual weight that is applied to the connection cost of a client depends highly on the (random) connection costs of the other clients.

We use four key ingredients to overcome this technical hurdle.

First, we show that the algorithm provides a constant-factor approximation for rectangular weight vectors. This already answers the open problem stated in [2, 3]. In our analysis, the connection cost of a single client is partly charged to a *deterministic budget* related to a *combinatorial bound* based on guessing, and partly to a *probabilistic budget* whose expected value is bounded with respect to the fractional LP-solution. This approach allows to limit the above-described problematic effect of the variance of individual client connection costs on the value of the ordered objective function of ORDERED k -MEDIAN.

Second, we show a surprising *modularity* of Charikar and Li's rounding process. The solution computed by this process can be related to the above-mentioned combinatorial and fractional bounds simultaneously with respect to *all* rectangular objectives. This property is oblivious to the objective with respect to which the input fractional solution was optimized.

Third, we *decompose* an arbitrary non-increasing weight vector into a convex combination of rectangular objectives. The aforementioned modularity property provides a bound for each of those objectives. We show that those bounds nicely combine to a global bound on the approximation ratio giving a constant-factor approximation with respect to a combinatorial bound and a fractional bound both under the original, general weight objective.

A straightforward application of this approach incorporating weight bucketing gives only quasi-polynomial time due to the guessing part. To achieve a truly polynomial time algorithm we apply a clever distance bucketing approach by Aouad and Segev [3], which guesses for each distance bucket the average weight applied to this bucket by some optimal solution. Our analysis approach applies also to this more intricate setting but turns out technically more involved.

2 Definitions

Definition 2.1. *In the METRIC ORDERED k -MEDIAN problem we are given: a finite set of facilities \mathcal{F} , a set of clients \mathcal{C} , $|\mathcal{C}| = n$, a metric cost function $c: \mathcal{F} \times \mathcal{C} \rightarrow \mathbb{R}_{\geq 0}$, an integer $k \geq 1$ as the number of facilities to open and a non-increasing weight vector $w = (w_1, \dots, w_n)$. For a subset $\mathcal{W} \subseteq \mathcal{F}$ and client $j \in \mathcal{C}$, we define $c_j(\mathcal{W}) = \min_{i \in \mathcal{W}} c(i, j)$ as the smallest connection cost of j to a facility in \mathcal{W} . We sort the values $c_j(\mathcal{W}), j \in \mathcal{C}$ in non-increasing order i.e. we define*

$c^\rightarrow(\mathcal{W}) = (c_j^\rightarrow(\mathcal{W}) : 1 \leq j \leq n)$ such that $\{c_j^\rightarrow(\mathcal{W}) : 1 \leq j \leq n\} = \{c_j(\mathcal{W}) : j \in \mathcal{C}\}$ and $c_j^\rightarrow(\mathcal{W}) \geq c_{j'}^\rightarrow(\mathcal{W})$ whenever $1 \leq j \leq j' \leq n$. The connection cost of \mathcal{W} is the weighted sum $\text{cost}(\mathcal{W}) = \sum_{j=1}^n w_j \cdot c_j^\rightarrow(\mathcal{W})$. The goal is to find a set $\mathcal{W} \subseteq \mathcal{F}, |\mathcal{W}| = k$ that minimizes the connection cost.

In the rest of the paper we say ORDERED k -MEDIAN for METRIC ORDERED k -MEDIAN because non-metric cost function does not allow us to obtain any non-trivial approximation (unless $\text{P} = \text{NP}$). In what follows we will assume w.l.o.g. that $w_1 = 1$ in the above definition. Let $j \in \mathcal{C}$ be a client. Then $\mathcal{B}(j, r)$ denotes the set of all facilities $i \in \mathcal{F}$ with $c_{ij} < r$, that is, $\mathcal{B}(j, r)$ is an open ball (in the set of facilities) of radius r around j .

Definition 2.2. Consider an instance of ORDERED k -MEDIAN. A reduced cost function c^r is a (not necessarily metric) function $c^r : \mathcal{F} \times \mathcal{C} \rightarrow \mathbb{R}_{\geq 0}$ such that for all $i, i' \in \mathcal{F}$ and $j, j' \in \mathcal{C}$ we have that $c_{ij}^r \leq c_{ij}$ and that $c_{ij} \leq c_{i'j'}$ implies $c_{ij}^r \leq c_{i'j'}^r$.

Reduced cost functions arise naturally for ORDERED k -MEDIAN since in its objective function non-increasingly sorted distances are multiplied by non-increasing weights ≤ 1 .

Definition 2.3. RECTANGULAR ORDERED k -MEDIAN is a special case of ORDERED k -MEDIAN problem with weights $w_1 = w_2 = \dots = w_\ell = 1$ and $w_{\ell+1} = w_{\ell+2} = \dots = w_n = 0$ for some $\ell \in [n]$. For any $\mathcal{W} \subseteq \mathcal{F}$ let $\text{cost}_\ell(\mathcal{W})$ denote the objective function of \mathcal{W} for this problem.

Note that RECTANGULAR ORDERED k -MEDIAN with $\ell = 1$ is equivalent to k -CENTER and RECTANGULAR ORDERED k -MEDIAN with $\ell = n$ is equivalent to k -MEDIAN. In what follows, missing proofs can be found in the appendix.

3 Algorithmic Framework

Our algorithms consist of two parts: An *LP-solving* and an *LP-rounding* part.

In the *LP-solving* part, we compute an optimal solution to an LP-relaxation, which is (apart from the objective function) identical to the standard LP relaxation for k -MEDIAN. However, instead of using the input metrics c in the objective function, we employ a reduced cost function c^r . Intuitively in c^r the distances are multiplied by roughly the same weights as in a guessed optimal solution.

In the *LP-rounding* part the fractional solution provided by the above-described guessing will be rounded to an integral solution by applying the algorithm of Charikar and Li [14]. In contrast to the LP-solving part, this algorithm operates, however, in the *original* metric space rather than in the (generally non-metric) reduced cost space.

3.1 LP-Relaxation

Let $\text{LP}(c^r)$ be the following relaxation of a natural ILP formulation of k -MEDIAN under some reduced cost function c^r .

$$\text{minimize} \quad \sum_{i \in \mathcal{F}, j \in \mathcal{C}} c_{ij}^r x_{ij} \quad \text{s.t.} \quad (1)$$

$$x_{ij} \leq y_i \quad i \in \mathcal{F}, j \in \mathcal{C} \quad (2)$$

$$\sum_{i \in \mathcal{F}} x_{ij} = 1 \quad j \in \mathcal{C} \quad (3)$$

$$\sum_{i \in \mathcal{F}} y_i = k \quad (4)$$

$$0 \leq x_{ij}, y_i \leq 1 \quad i \in \mathcal{F}, j \in \mathcal{C} \quad (5)$$

Here, y_i denotes how much facility i is open (0—closed, 1—opened) and x_{ij} indicates how much client j is served by facility i (0—non-served, 1—served). Equality (4) ensures that exactly k facilities are opened (possibly fractionally), (3) guarantee that each client is served (possibly fractionally). (2) do not allow a facility to serve a client more than how much it is opened. For each client $j \in \mathcal{C}$ let $c_{\text{av}}^r(j) = \sum_{i \in \mathcal{F}} c_{ij}^r x_{ij}$ denote the fractional (or average) reduced connection cost of j .

3.2 Guessing and LP-Solving

Note that if $c^r = c$ where c is the input metrics, $\text{LP}(c)$ becomes the standard LP relaxation for the classical k -MEDIAN objective. In order to obtain a valid lower bound $\text{LP}(c^r)$ for a ORDERED k -MEDIAN instance, we employ guessing of certain distances in an optimal solution. The details of the guessing are setting-specific and are thus described later. We can w.l.o.g. assume that we compute an optimal solution (x, y) to $\text{LP}(c^r)$ such that for all $i \in \mathcal{F}, j \in \mathcal{C}$ we have that $x_{ij} \in \{0, y_i\}$ and that $y_i > 0$. Also, since c^r preserves the order of the distances, we can assume that if y is kept fixed then x optimizes $\sum_{i \in \mathcal{F}, j \in \mathcal{C}} c_{ij} x_{ij}$ and say that x be *distance-optimal*. See appendix for details.

3.3 LP-Rounding: Dependent Rounding Approach of Charikar and Li

We round the fractional solution obtained in the LP-solving phase to an integral solution by the (slightly modified) LP-rounding process of Charikar and Li [14] for k -MEDIAN.

To apply this algorithm note that the feasibility of a solution (x, y) to $\text{LP}(c^r)$ does *not* depend on the cost vector c^r . This enables us to compute an optimum solution (x, y) to $\text{LP}(c^r)$ for some appropriate reduced cost function and to subsequently apply the rounding process of Charikar and Li (which operates on the original metrics c) to the solution (x, y) . In the analysis, we have to exploit how c^r and c are related in order to bound the approximation ratio of the algorithm.

We now describe the rounding algorithm of Charikar and Li, which consists of four phases: a clustering phase, a bundling phase, a matching phase, and a sampling phase (see Algorithm 1). Below we give some intuition on the different phases. More formal arguments will be given later.

The purpose of the *clustering procedure* is to compute a set $\mathcal{C}' \subseteq \mathcal{C}$ of *cluster centers* so that each client $j \in \mathcal{C}$ is “close” to some cluster center $j' \in \mathcal{C}'$ and so that the cluster centers are “far” from each other. We thus may think of the cluster centers representing all remaining clients. The implementation of the procedure and the meaning of “close” and “far” is application-specific and will thus be described later.

In the *bundling phase* each cluster center $j \in \mathcal{C}'$ is associated with a bundle \mathcal{U}_j of facilities of total fractional opening at least $1/2$, so that we can show that the bundles are pairwise disjoint².

In the *matching phase* cluster centers are paired in a greedy manner. As we will show that the volume of each bundle is at least $1/2$ the total volume of the bundles of a matched pair is at least 1. This will ensure that in the subsequent *sampling phase* at least one facility per pair is opened.

In the sampling phase we use the dependent randomized rounding procedure described by Charikar and Li [14] to open facilities and obtain a feasible solution. We do not describe the details

²Our version of the algorithm is actually slightly simpler here than the original one, which is sufficient for constant-factor approximations. In the original versions the bundles are based on larger balls that may overlap but are made disjoint in a greedy manner. This allows Charikar and Li to obtain an improved approximation factor.

Algorithm 1: Rounding Algorithm by Charikar and Li [14]

Data: feasible fractional solution (x, y) to $LP(c^r)$ satisfying the properties of Lemma A.2
Result: set of k facilities

```
/* Clustering phase */
/* run a clustering procedure to compute a set  $C' \subseteq C$  of cluster centers so that
each client  $j \in C$  is ‘‘close’’ to some cluster center  $j' \in C'$  and so that the cluster
centers are ‘‘far’’ from each other */
1  $C' \leftarrow \text{Clustering}(x, y);$ 
/* Bundling phase */
2 for  $j \in C'$  do
3    $R_j \leftarrow \frac{1}{2} \min_{j' \in C', j' \neq j} (c_{jj'});$ 
4    $\mathcal{U}_j \leftarrow \{i \in \mathcal{F} : i \in \mathcal{B}(j, R_j) \text{ and } x_{ij} > 0\};$ 
/* Matching phase */
5  $\mathcal{M} \leftarrow \emptyset;$ 
6 while there are unmatched clients in  $C'$  do
7    $\mathcal{M} \leftarrow \mathcal{M} \cup \{j, j'\};$ 
/* Sampling phase (dependent rounding) */
/* Apply dependent randomized rounding as described by Charikar and Li [14]
preserving the marginals for the individual facilities, bundles, matched pairs in
 $\mathcal{M}$ , and set  $\mathcal{F}$  */
8 return  $\text{DependentRounding}(x, y, \{U_j\}_j, \mathcal{M}, \mathcal{F})$ 
```

of the implementation here but use it as a ‘‘black box’’ satisfying the following properties (as in the original work of Charikar and Li):

Lemma 3.1. *The procedure DependentRounding in Algorithm 1 can be implemented such that the following holds for any input (x, y) being a feasible solution to $LP(c^r)$.*

- (i) *Each facility $i \in \mathcal{F}$ is opened with probability precisely y_i ,*
- (ii) *in each bundle \mathcal{U}_j with $j \in C'$ a facility is opened with probability precisely $\text{vol}(\mathcal{U}_j)$,*
- (iii) *for each matched pair (j, j') in \mathcal{M} at least one facility in $\mathcal{U}_j \cup \mathcal{U}_{j'}$ will be opened,*
- (iv) *in total at most k facilities are opened.*

4 Rectangular Weight Vectors

Theorem 4.1. *There exists a polynomial time randomized algorithm for RECTANGULAR ORDERED k -MEDIAN that gives a 15-approximation in expectation.*

To proof this theorem, we need to fill in the following two missing parts of the framework: Guessing of the reduced cost space and the clustering procedure in the rounding part.

4.1 Guessing and Reduced Costs

In the LP-solving phase, we guess the value T of ℓ -th largest distance in an optimum solution to RECTANGULAR ORDERED k -MEDIAN. (This is the smallest distance that is counted in the total connection cost with non-zero weight.) As the correct guess of T is the distance between a client and a facility the guessing can be performed by considering only $\mathcal{O}(mn)$ options for T .

For each $i \in \mathcal{F}, j \in \mathcal{C}$, we define the *reduced cost*

$$c_{ij}^T = \begin{cases} c_{ij} & \text{if } c_{ij} \geq T, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

that will be used as a cost function in our LP for the ORDERED k -MEDIAN.

An optimal solution (x, y) to $\text{LP}(c^T)$ is a feasible solution for $\text{LP}(c)$ as well. As introduced in Section 3.1, we use $c_{\text{av}}(j) = \sum_{i \in \mathcal{F}} x_{ij} \cdot c_{ij}$ and $c_{\text{av}}^T(j) = \sum_{i \in \mathcal{F}} x_{ij} \cdot c_{ij}^T$ to denote the average connection cost and the average reduced connection cost of a client $j \in \mathcal{C}$, respectively.

4.2 Dedicated Clustering

The following two clustering methods will be considered. We first analyze using Algorithm 2.

Algorithm 2: DedicatedClustering	Algorithm 3: ObliviousClustering
Data: feasible fractional solution (x, y) to $\text{LP}(c)$	Data: feasible fractional solution (x, y) to $\text{LP}(c)$
Result: set $\mathcal{C}' \subseteq \mathcal{C}$ of cluster centers	Result: set $\mathcal{C}' \subseteq \mathcal{C}$ of cluster centers
1 $\mathcal{C}' \leftarrow \emptyset;$	1 $\mathcal{C}' \leftarrow \emptyset;$
2 $\mathcal{C}'' \leftarrow \mathcal{C};$	2 $\mathcal{C}'' \leftarrow \mathcal{C};$
3 $\underline{c}_{\text{av}}^T(j) \leftarrow \sum_{i \in \mathcal{F}} x_{ij} \cdot \underline{c}_{ij}^T$ for all $j \in \mathcal{C};$ DIFF	3 $\underline{c}_{\text{av}}(j) \leftarrow \sum_{i \in \mathcal{F}} x_{ij} \cdot \underline{c}_{ij}$ for all $j \in \mathcal{C};$
4 while \mathcal{C}'' is non empty do	4 while \mathcal{C}'' is non empty do
5 take $j \in \mathcal{C}''$ with the smallest $\underline{c}_{\text{av}}^T(j);$ DIFF	5 take $j \in \mathcal{C}''$ with the smallest $\underline{c}_{\text{av}}(j);$
6 add j to $\mathcal{C}';$	6 add j to $\mathcal{C}';$
7 delete from \mathcal{C}'' client $j;$	7 delete from \mathcal{C}'' client $j;$
8 delete from \mathcal{C}'' all clients j' with	8 delete from \mathcal{C}'' all clients j' with
$c_{jj'} \leq \underline{4c}_{\text{av}}^T(j') + 4T$ DIFF	$c_{jj'} \leq \underline{4c}_{\text{av}}(j')$
9 return \mathcal{C}'	9 return \mathcal{C}'

This clustering procedure is very similar to the one of Charikar and Li (see also Section 4.4 below) except for the fact that the procedure needs to know the threshold T of the guessing phase. (Note that we use T explicitly but also implicitly in the average reduced cost $c_{\text{av}}^T(j)$.) This dependence allows a simpler and better analysis for RECTANGULAR ORDERED k -MEDIAN. In Section 4.4, we will describe how to get rid of this dependency, which allows us to generalize the result.

4.3 Analysis of the Algorithm

In the following we analyze Algorithm 1 using the procedure DedicatedClustering.

Observation 4.2. We have $c_{ij}^T \leq c_{ij} \leq c_{ij}^T + T$ for any $i \in \mathcal{F}, j \in \mathcal{C}$ and consequently $c_{\text{av}}^T(j) \leq c_{\text{av}}(j) \leq c_{\text{av}}^T(j) + T$.

Lemma 4.3. The following two statements are true

- (i) For any $j, j' \in \mathcal{C}'$ we have that $c_{jj'} > 4 \max(c_{\text{av}}^T(j), c_{\text{av}}^T(j')) + 4T$.
- (ii) For any $j \in \mathcal{C} \setminus \mathcal{C}'$ there is a client $j' \in \mathcal{C}'$ with $c_{\text{av}}^T(j') \leq c_{\text{av}}^T(j)$ and $c_{jj'} \leq 4c_{\text{av}}^T(j) + 4T$.

Lemma 4.4. The following two statements are true

- (i) $\text{vol}(\mathcal{U}_j) \geq 0.5$ for all $j \in \mathcal{C}'$
- (ii) $\mathcal{U}_j \cap \mathcal{U}_{j'} = \emptyset$ for all $j, j' \in \mathcal{C}', j \neq j'$

Proof of Theorem 4.1. Let OPT , OPT^* and ALG be the values of an optimal solution, $\text{LP}(c_r)$ and Algorithm 1, respectively. Note that $\text{OPT}^* \leq \text{OPT}$.

The idea of the proof is to provide for each client an upper bound on the distance C_j (according to the original distance c) traveled by this client. The upper bound is paid for by two budgets D_j and X_j . The “deterministic” budget D_j is $5T$. The “probabilistic” budget X_j is a random variable (depending on the random choices made by the algorithm).

We will show below that (by a suitable choice of X_j) the connection cost C_j of j can actually be upper bounded by $D_j + X_j$ and that $\mathbb{E}[X_j] \leq 10 \cdot c_{\text{av}}^T(j)$. We claim that this will complete our proof of a 15-approximation. To this end, note that at most ℓ clients j will pay the deterministic budget $D_j = 5T$ because at most ℓ distances are actually accounted for in the objective function. Unfortunately, an analogous reasoning does not hold true for the expected value of the random variables X_j . (For example, note that $\mathbb{E}[\max(X_1, \dots, X_n)]$ is generally unbounded in $\max(\mathbb{E}[X_1], \dots, \mathbb{E}[X_n])$ in the case of $\ell = 1$.) However, we can just sum over *all* those random variables obtaining an upper bound on the total expected connection cost:

$$\mathbb{E}[\text{ALG}] \leq D_j \cdot \ell + \sum_{j \in C} \mathbb{E}[X_j] \leq 5\ell \cdot T + 10 \cdot \sum_{j \in C} c_{\text{av}}^T(j) \leq 15 \cdot \text{OPT}.$$

For the last inequality note that by our guess of T we have that $\text{OPT} \geq \ell \cdot T$ and from the definition of $\text{LP}(c_r)$ we have $\text{OPT}^* = \sum_{j \in C} c_{\text{av}}^T(j)$. To show that $C_j \leq D_j + X_j$ consider an arbitrary client j with connection cost C_j . We incrementally construct our upper bound on C_j starting with 0. Each increment will be either charged to D_j or X_j .

Consider a client j and the cluster center j' it is assigned to (possibly $j = j'$). We have that $c_{jj'} \leq 4c_{\text{av}}^T(j) + 4T$ by Lemma 4.3 (ii). We charge $4T$ to D_j and $4c_{\text{av}}^T(j)$ with probability 1 to X_j .

We now describe how to pay for the transport from j' to an open facility. There are two cases to distinguish. Either a facility within a radius T around j' is opened or not. If yes, then this cost can be covered by charging an additional amount of T to D_j . In this case the total cost is upper bounded by $D_j = 5T$ and $\mathbb{E}[X_j] = 4c_{\text{av}}^T(j)$.

If no facility within a radius T around j' is opened then observe that for all facility i with $c_{ij'} \geq T$ we have that $c_{ij'}^T = c_{ij'}$. We now continue to bound the connection cost for this case. Let j'' be the closest client distinct from j' in \mathcal{C}' . We consider the case where j' and j'' are not matched. (The case where they are matched is simpler.) Let j''' be the client in \mathcal{C}' to which j'' is matched, i.e., $(j'', j''') \in \mathcal{M}$. By the dependent rounding process one facility in $\mathcal{U}_{j''} \cup \mathcal{U}_{j'''}$ will be opened. We have that $c_{j'j''} = 2R_{j'} = 2R$ and thus $c_{j''j'''} \leq 2R$ and $R_{j''}, R_{j'''} \leq R$ (otherwise, j'' would not have been matched with j''' but with j').

This means that, in case no facility is opened in the bundle $\mathcal{U}_{j'}$ the client j travels an additional distance of at most $\max(c_{j'j''} + R_{j''}, c_{j'j''} + c_{j''j'''} + R_{j'''}) \leq 2R + 2R + R \leq 5R$.

If a facility is opened in the bundle $\mathcal{U}_{j'}$ then we charge this additional connection cost to X_j . The contribution of this case to $\mathbb{E}[X_j]$ is (by Properties (i) and (ii) of Lemma 3.1) at most

$$\sum_{i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', T)} y_i c_{ij'} = \sum_{i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', T)} x_{ij'} c_{ij'} = \sum_{i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', T)} x_{ij'} c_{ij'}^T \leq \sum_{i \in \mathcal{F}_{j'}} x_{ij'} c_{ij'}^T = c_{\text{av}}^T(j').$$

Here, the first equality follows by Property (ii) of Lemma A.2. The second equality follows because we assume that no facility is opened in $\mathcal{B}(j', T)$ and since $c_{ij'} = c_{ij'}^T$ for all $i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', T)$.

We finally handle the case where no facility in $\mathcal{U}_{j'}$ is opened and where j additionally travels a distance of at most $5R$. We charge this additional cost to X_j . We bound the probability that this case occurs. We claim that $\text{vol}(\mathcal{U}_{j'})$ is at least $1 - c_{\text{av}}^T(j')/R$. To see this, recall that $2R \geq c_{j''j'''} > 4 \max(c_{\text{av}}^T(j''), c_{\text{av}}^T(j''')) + 4T$ thus $R > T$. Note that the reason of adding the quantity

$4T$ in the clustering phase (line 8) is to have the property $R > T$ (in the original algorithm of Charikar-Li [14] this property is not necessarily satisfied). Using this, for all facilities in $\mathcal{F}_{j'} \setminus \mathcal{U}_{j'}$ we have that $c_{ij'}^T = c_{ij}$ because $R > T$. Hence

$$\begin{aligned} c_{av}^T(j') &\geq \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{B}(j', T)} x_{ij'} \cdot c_{ij'}^T = \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{B}(j', T)} x_{ij'} \cdot c_{ij'} \geq \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{U}_{j'}} x_{ij'} \cdot c_{ij'} \\ &\geq R \cdot \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{U}_{j'}} x_{ij'} = R \cdot \text{vol}(\mathcal{F}_{j'} \setminus \mathcal{U}_{j'}) = R \cdot (1 - \text{vol}(\mathcal{U}_{j'})), \end{aligned}$$

which implies the claim. Here, note that $\mathcal{B}(j', T) \subseteq \mathcal{U}_{j'}$ because $R > T$. This means that j travels the additional distance of $5R$ with probability at most $c_{av}^T(j')/R$ and hence the contribution to $\mathbb{E}[X_j]$ is upper bounded by $5 \cdot c_{av}^T(j')$. Summarizing, for the case when no facility is opened within $\mathcal{B}(j', T)$ we can upper bound $\mathbb{E}[X_j]$ by:

- a cost of serving client j through the closest cluster center j' that is $4 \cdot c_{av}^T(j)$, plus
- a value $c_{av}^T(j')$ for the case when a facility is opened within bundle $\mathcal{U}_{j'}$, plus
- a value $5R$ with probability at most $c_{av}^T(j')/R$ when no facility is opened within $\mathcal{U}_{j'}$.

Hence $\mathbb{E}[X_j] \leq 4 \cdot c_{av}^T(j) \cdot 1 + c_{av}^T(j') \cdot 1 + 5R \cdot c_{av}^T(j')/R \leq 10 \cdot c_{av}^T(j)$, by taking into account that $c_{av}^T(j') \leq c_{av}^T(j)$. Moreover we charged again at most $5T$ to D_j in this case. In the end we have the desired two upper bounds for both budgets for completing the proof: $D_j \leq 5T$, $\mathbb{E}[X_j] \leq 10 \cdot c_{av}^T(j)$. \square

4.4 Oblivious Clustering

In Algorithm 1, we are working on the original metrics c but still DedicatedClustering described in the previous section depends on our guessed parameter T and the reduced metrics c^T . In this section, we show that we can apply the original clustering of Charikar and Li that works solely on the input metrics c and that is thus *oblivious* of the guessing phase. In particular, we use the Oblivious Clustering procedure as described in Algorithm 3.

Using Oblivious Clustering, we can prove the following version of Theorem 4.1. While the constants proven in the following lemma are weaker than the ones for DedicatedClustering, it exhibits a surprising modularity that is a key ingredient to later handle the general case. In particular, the clustering (and thus the whole rounding phase) are unaware (oblivious) of the cost vector \bar{c} with respect to which we optimized $\text{LP}(\bar{c})$. Secondly, the bound proven in the lemma holds for *any* rectangular objective function of ORDERED k -MEDIAN (specified by parameter ℓ), threshold T and the corresponding average reduced cost and may be unrelated to the cost function \bar{c} that we optimized to obtain the fractional solution (x, y) .

Lemma 4.5. *Consider a feasible fractional solution (x, y) to $\text{LP}(c)$ where x is distance-optimal. Let $\ell \geq 1$ be a positive integer, let $T \geq 0$ be arbitrary. Then we have $\mathbb{E}[\text{cost}_\ell(A)] \leq 19\ell T + 19 \sum_{j \in C} c_{av}^T(j)$ where A is the (random) solution output by the Algorithm 1 with oblivious clustering.*

By Lemma 4.5 we obtain that our algorithm with oblivious clustering yields a 38-approximation.

5 Handling the General Case

Consider an arbitrary instance of ORDERED k -MEDIAN. Let w be the weight vector and let \bar{w} the sorted weight vector using the same weights as w but without repetition. Let R be the number of distinct weight in both weight vectors. W.l.o.g. we assume that all distances c_{ij} for $i \in \mathcal{F}, j \in \mathcal{C}$ are pairwise distinct. (This can be achieved by slightly perturbing the input distances.) To apply our algorithmic framework, we guess *thresholds* T_r for $r = 1, \dots, R$ such that T_r is the smallest distance c_{ij} that is multiplied by weight of value \bar{w}_r in some fixed optimum solution. To guess the thresholds T_r we check $(nm)^R$ many candidates. Additionally, we define $T_0 = \infty$. We have $T_r < T_{r-1}$ for $r = 1, \dots, R$ because we assumed pairwise distinct distances. For each $i \in \mathcal{F}, j \in \mathcal{C}$ we assign the connection cost c_{ij} to the weight $w(i, j) = \bar{w}_r$, where $T_r \leq c_{ij} < T_{r-1}$. This leads us to the following definition of our reduced cost function $c_{ij}^r = c_{ij} \cdot w(i, j)$ for all $i \in \mathcal{F}, j \in \mathcal{C}$. We compute an optimal solution (x, y) to $LP(c^r)$ and apply Algorithm 1 to (x, y) .

Lemma 5.1. *The above-described algorithm is a 38-approximation algorithm for ORDERED k -MEDIAN that makes $\mathcal{O}((nm)^R)$ many calls to Algorithm 1 with oblivious clustering, where R is the number of distinct weights in the weight vector w .*

Proof. Let $A \subseteq \mathcal{F}$ be the (random) solution output by the algorithm. Let OPT be the optimum objective function. For each $r = 1, \dots, R$ let ℓ_r be the largest index such that $w_{\ell_r} = \bar{w}_r$. From Lemma 4.5 we have for all $r = 1, \dots, R$

$$\mathbb{E}[\text{cost}_{\ell_r}(A)] \leq 19 \cdot \ell_r T_r + 19 \cdot \sum_{j \in \mathcal{C}} c_{\text{av}}^{T_r}(j). \quad (7)$$

We decompose $\text{cost}(A)$ into rectangular “pieces” (defining $w_{R+1} = 0$)

$$\begin{aligned} \mathbb{E}[\text{cost}(A)] &= \mathbb{E}\left[\sum_{\ell=1}^n w_{\ell} \cdot c_{\ell}^{\rightarrow}(A)\right] = \mathbb{E}\left[\sum_{r=1}^R \sum_{s=1}^{\ell_r} (\bar{w}_r - \bar{w}_{r+1}) \cdot c_s^{\rightarrow}(A)\right] = \mathbb{E}\left[\sum_{r=1}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot \text{cost}_{\ell_r}(A)\right] \\ &= \sum_{r=1}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot \mathbb{E}[\text{cost}_{\ell_r}(A)] \stackrel{(7)}{\leq} 19 \cdot \sum_{r=1}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot \ell_r T_r + 19 \cdot \sum_{r=1}^R \sum_{j \in \mathcal{C}} (\bar{w}_r - \bar{w}_{r+1}) \cdot c_{\text{av}}^{T_r}(j). \end{aligned} \quad (8)$$

We will bound this in terms of OPT . We know that an optimal solution pays at least cost T_r for weight in w equal to \bar{w}_r for $r = 1, \dots, R$. Therefore, defining $\ell_0 = 0$ and $\bar{w}_{R+1} = 0$ we have

$$\begin{aligned} \text{OPT} &\geq \sum_{r=1}^R \bar{w}_r \cdot (\ell_r - \ell_{r-1}) T_r = \sum_{r=1}^R \bar{w}_r \cdot \ell_r T_r - \sum_{r=2}^R \bar{w}_r \cdot \ell_{r-1} T_r \geq \\ &\sum_{r=1}^R \bar{w}_r \cdot \ell_r T_r - \sum_{r=2}^R \bar{w}_r \cdot \ell_{r-1} T_{r-1} = \sum_{r=1}^R \bar{w}_r \cdot \ell_r T_r - \sum_{r=1}^R \bar{w}_{r+1} \cdot \ell_r T_r = \sum_{r=1}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot \ell_r T_r. \end{aligned} \quad (9)$$

Moreover, we have

$$\begin{aligned} \sum_{r=1}^R \sum_{j \in \mathcal{C}} (\bar{w}_r - \bar{w}_{r+1}) \cdot c_{\text{av}}^{T_r}(j) &= \sum_{r=1}^R \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} (\bar{w}_r - \bar{w}_{r+1}) \cdot x_{ij} \cdot c_{ij}^{T_r} = \\ &\sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot \sum_{r=1}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot c_{ij}^{T_r} = \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot \sum_{r: \bar{w}_r \leq w(i, j)}^R (\bar{w}_r - \bar{w}_{r+1}) \cdot c_{ij} = \end{aligned}$$

$$\sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot w(i, j) \cdot c_{ij} = \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot c_{ij}^r \stackrel{(1)}{\leq} \text{OPT}. \quad (10)$$

Thus, we finally have $\mathbb{E}[\text{cost}(A)] \stackrel{(8),(9),(10)}{\leq} 38 \cdot \text{OPT}$. \square

An immediate consequence of the lemma is a constant-factor approximation algorithm for ORDERED k -MEDIAN with a constant number of different weights.

Using standard bucketing arguments and neglecting sufficiently small weights, we can “round” an arbitrary weight vector into a weight vector with only a logarithmic number of different weights losing a factor of $1 + \epsilon$ in approximation. Plugging this into Lemma 4.5, we can obtain a $(38 + \epsilon)$ -approximation algorithm for the general case in time $(nm)^{\mathcal{O}(1/\epsilon \log(n))}$.

Achieving Polynomial Time To obtain a truly-polynomial time algorithm we use the clever bucketing approach proposed by Aouad et al. [3]. In this approach the *distances* are grouped into logarithmically many distance classes thereby losing a factor $1 + \epsilon$. For each distance class the *average* weight is guessed up to a factor of $1 + \epsilon$. The crucial point is, that this guessing can be achieved in polynomial time because the average weights are non-decreasing with increasing distance class. This leads to a reduced cost function based on average weights. The analysis of the resulting analysis decomposes the weight vector into $n = |\mathcal{C}|$ many rectangular objectives. While the proof strategy is similar in spirit to the one of Lemma 5.1 it turns out to be technically more involved (see appendix for details).

Theorem 5.2. *For any $\epsilon > 0$, there is an $(38 + \epsilon)$ -approximation algorithm for ORDERED k -MEDIAN with running time $(nm)^{\mathcal{O}(1/\epsilon \log(1/\epsilon))}$.*

6 Concluding Remarks and Open Questions

We have obtained a constant-factor approximation for ORDERED k -MEDIAN. We have extended the less detailed version of the analysis of the algorithm by Charikar and Li [14] for k -MEDIAN, and hence our constants can probably be improved. It would be interesting to see, if our ideas can be used for other problems with ordered objectives.

Acknowledgments. J. Byrka and J. Spoerhase were supported by the NCN grant number 2015/18/E/ST6/00456. K. Sornat was supported by the NCN grant number 2015/17/N/ST6/03684.

References

- [1] Sara Ahmadian, Ashkan Norouzi-Fard, Ola Svensson, and Justin Ward. Better guarantees for k -means and Euclidean k -median by primal-dual algorithms. In *Proc. 58th IEEE Symposium on Foundations of Computer Science (FOCS'17)*, 2017.
- [2] Soroush Alamdari and David Shmoys. A bicriteria approximation algorithm for the k -center and k -median problems. In *Proceedings of the 15th International Workshop Approximation and Online Algorithms (WAOA'17)*, 2017.

- [3] Ali Aouad and Danny Segev. The ordered k -median problem: Surrogate models and approximation algorithms. Under review, available at <http://www.mit.edu/~aaouad/MOR-ordered-median.pdf> [2018-02-20].
- [4] Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. Local search heuristics for k -median and facility location problems. *SIAM Journal on computing*, 33(3):544–562, 2004.
- [5] Yair Bartal. On approximating arbitrary metrics by tree metrics. In *Proc. Thirtieth Annual ACM Symposium on the Theory of Computing (STOC’98)*, pages 161–168, 1998.
- [6] Dimitris Bertsimas and Rahul Mazumder. Least quantile regression via modern optimization. *Ann. Statist.*, 42(6):2494–2525, 12 2014.
- [7] Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization and network flows. *Math. Program.*, 98(1-3):49–71, 2003.
- [8] Dimitris Bertsimas and Robert Weismantel. *Optimization over integers*. Athena Scientific, 2005.
- [9] Natasha Boland, Patricia Domínguez-Marín, Stefan Nickel, and Justo Puerto. Exact procedures for solving the discrete ordered median problem. *Computers & OR*, 33(11):3270–3300, 2006.
- [10] Paul S. Bradley, Usama M. Fayyad, and Olvi L. Mangasarian. Mathematical programming for data mining: Formulations and challenges. *INFORMS Journal on Computing*, 11(3):217–238, 1999.
- [11] Jarosław Byrka, Thomas Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k -median and positive correlation in budgeted optimization. *ACM Trans. Algorithms*, 13(2):23:1–23:31, 2017.
- [12] Deeparnab Chakrabarty and Chaitanya Swamy. Interpolating between k -median and k -center: Approximation algorithms for ordered k -median. *CoRR*, abs/1711.08715, 2017.
- [13] Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the k -median problem. In *Proc. 31st Annual ACM Symposium on Theory of Computing (STOC’99)*, pages 1–10, 1999.
- [14] Moses Charikar and Shi Li. A dependent LP-rounding approach for the k -median problem. In *Proc. 39th International Colloquium on Automata, Languages, and Programming (ICALP’12)*, pages 194–205, 2012.
- [15] Patricia Domínguez-Marín, Stefan Nickel, Pierre Hansen, and Nenad Mladenovic. Heuristic procedures for solving the discrete ordered median problem. *Annals OR*, 136(1):145–173, 2005.
- [16] Zvi Drezner and Stefan Nickel. Constructing a DC decomposition for ordered median problems. *J. Global Optimization*, 45(2):187–201, 2009.
- [17] Zvi Drezner and Stefan Nickel. Solving the ordered one-median problem in the plane. *European Journal of Operational Research*, 195(1):46–61, 2009.
- [18] Inmaculada Espejo, Antonio M. Rodríguez-Chía, and C. Valero. Convex ordered median problem with ℓ_p -norms. *Computers & OR*, 36(7):2250–2262, 2009.

- [19] Elena Fernández, Miguel A. Pozo, Justo Puerto, and Andrea Scozzari. Ordered weighted average optimization in multiobjective spanning tree problem. *European Journal of Operational Research*, 260(3):886–903, 2017.
- [20] David G. Harris, Thomas Pensyl, Aravind Srinivasan, and Khoa Trinh. Symmetric randomized dependent rounding. *CoRR*, abs/1709.06995, 2017.
- [21] Dorit S. Hochbaum and David B. Shmoys. A best possible heuristic for the k -center problem. *Mathematics of Operations Research*, 10(2):180–184, 1985.
- [22] Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, and Vijay V Vazirani. Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. *Journal of the ACM (JACM)*, 50(6):795–824, 2003.
- [23] Jörg Kalcsics, Stefan Nickel, and Justo Puerto. Multifacility ordered median problems on networks: A further analysis. *Networks*, 41(1):1–12, 2003.
- [24] Jörg Kalcsics, Stefan Nickel, Justo Puerto, and Arie Tamir. Algorithmic results for ordered median problems. *Oper. Res. Lett.*, 30(3):149–158, 2002.
- [25] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu. A local search approximation algorithm for k -means clustering. *Computational Geometry*, 28(2-3):89–112, 2004.
- [26] Martine Labbé, Diego Ponce, and Justo Puerto. A comparative study of formulations and solution methods for the discrete ordered p -median problem. *Computers & OR*, 78:230–242, 2017.
- [27] Gilbert Laporte, Stefan Nickel, and Francisco Saldanha da Gama, editors. *Location Science*. Springer, 2015.
- [28] Shi Li and Ola Svensson. Approximating k -median via pseudo-approximation. *SIAM Journal on Computing*, 45(2):530–547, 2016.
- [29] S. Nickel. Discrete ordered Weber problems. In *Proc. Operations Research (OR’01)*, pages 71–76. Springer, 2001.
- [30] Stefan Nickel and Justo Puerto. A unified approach to network location problems. *Networks*, 34(4):283–290, 1999.
- [31] Stefan Nickel and Justo Puerto. *Location Theory - A Unified Approach*. Springer, 2009.
- [32] J. Puerto and F. R. Fernández. Geometrical properties of the symmetrical single facility location problem. *Journal of Nonlinear and Convex Analysis*, 1(3):321–342, 2000.
- [33] Justo Puerto, Dionisio Pérez-Brito, and Carlos G. García-González. A modified variable neighborhood search for the discrete ordered median problem. *European Journal of Operational Research*, 234(1):61–76, 2014.
- [34] Justo Puerto, A. B. Ramos, and Antonio M. Rodríguez-Chía. A specialized branch & bound & cut for single-allocation ordered median hub location problems. *Discrete Applied Mathematics*, 161(16-17):2624–2646, 2013.

- [35] Justo Puerto and Antonio M. Rodríguez-Chía. On the exponential cardinality of FDS for the ordered p -median problem. *Oper. Res. Lett.*, 33(6):641–651, 2005.
- [36] Justo Puerto, Antonio M Rodríguez-Chía, and Arie Tamir. Minimax regret single-facility ordered median location problems on networks. *INFORMS Journal on Computing*, 21(1):77–87, 2009.
- [37] Justo Puerto, Antonio M. Rodríguez-Chía, and Arie Tamir. Revisiting k -sum optimization. *Math. Program.*, 165(2):579–604, 2017.
- [38] Justo Puerto and Arie Tamir. Locating tree-shaped facilities using the ordered median objective. *Math. Program.*, 102(2):313–338, 2005.
- [39] A. M. Rodríguez-Chía, J. Puerto, D. Pérez-Brito, and J. A. Moreno. The p -facility ordered median problem on networks. *Top*, 13(1):105–126, Jun 2005.
- [40] Zorica Stanimirovic, Jozef Kratica, and Djordje Dugosija. Genetic algorithms for solving the discrete ordered median problem. *European Journal of Operational Research*, 182(3):983–1001, 2007.
- [41] Arie Tamir. The k -centrum multi-facility location problem. *Discrete Applied Mathematics*, 109(3):293–307, 2001.
- [42] Arie Tamir, Dionisio Pérez-Brito, and José A. Moreno-Pérez. A polynomial algorithm for the p -centdian problem on a tree. *Networks*, 32(4):255–262, 1998.

A Missing Proofs for Section 3

Below, we describe some basic normalization steps for a feasible solution (x, y) to $LP(c^r)$.

Definition A.1. Let (x, y) be a feasible solution to $LP(c^r)$ where c^r is some reduced cost function. We call the assignment x of clients to facilities distance-optimal if x minimizes $\sum_{i \in \mathcal{F}, j \in \mathcal{C}} c_{ij} x_{ij}$ when y is kept fixed.

Lemma A.2. We can w.l.o.g. assume that an optimal solution (x, y) to $LP(c^r)$ for some reduced cost function c^r satisfies the following properties.

- (i) For any facility $i \in \mathcal{F}$ we have $y_i > 0$,
- (ii) for any $i \in \mathcal{F}, j \in \mathcal{C}$ we have $x_{ij} \in \{0, y_i\}$,
- (iii) the assignment x is distance-optimal.

Proof of Lemma A.2. To see the third property fix the opening vector y and some client j . Now sort all facilities i in non-decreasing order of their distance c_{ij} to j and greedily assign as much of the remaining demand of j to the current facility i (respecting the constraint $x_{ij} \leq y_i$). Stop when the full demand of j is served and repeat this process for all clients. Since the reduced cost function c^r respects the order of the original distances (see definition) the resulting assignment is optimal also under the reduced cost function.

The first and second properties are folklore and can be achieved by removing or duplicating facilities (see [14]). □

Proof of Lemma 3.1. Charikar and Li [14] describe how to implement the procedure satisfying the above claims. In doing so, the only requirement is that (x, y) be a (not necessarily optimal) feasible solution to $\text{LP}(c)$, that the volume of each bundle \mathcal{U}_j has volume at least $1/2$ and that the union of set families $\{\{y_i\}_{i \in \mathcal{F}}, \{U_j\}_{j \in \mathcal{C}'}, \{\mathcal{U}_j \cup \mathcal{U}_{j'} \mid (j, j') \in \mathcal{M}\}, \{\mathcal{F}\}$ form a laminar family. The latter claim on laminarity follows immediately from the construction in the algorithm. The validity of $\text{vol}(\mathcal{U}_j) \geq 1/2$ for $j \in \mathcal{C}'$ depends on the implementation of the clustering procedure and has thus to be proven for the specific implementation. \square

B Missing Proofs for Section 4

Proof of Lemma 4.3. To see (i) assume w.l.o.g. that j is considered before j' as a potential cluster center in the algorithm. Thus $c_{\text{av}}^T(j) \leq c_{\text{av}}^T(j')$. If $c_{jj'} \leq 4c_{\text{av}}^T(j') + 4T = 4 \max(c_{\text{av}}^T(j), c_{\text{av}}^T(j')) + 4T$ then j' would be deleted from \mathcal{C}'' when j is considered. A contradiction to the fact that j' is a cluster center.

In order to see (ii), consider an arbitrary client $j \in \mathcal{C} \setminus \mathcal{C}'$. As j is not a cluster center it was deleted from \mathcal{C}'' when some cluster center $j' \in \mathcal{C}'$ was considered. For this cluster center we have $c_{jj'} \leq 4c_{\text{av}}^T(j) + 4T$. \square

Proof of Lemma 4.4. To prove statement (i) consider an arbitrary $j \in \mathcal{C}'$. Let $j' \in \mathcal{C}'$ be such that $2R_j = c_{jj'}$. We have $c_{jj'} > 4c_{\text{av}}^T(j) + 4T \geq 4c_{\text{av}}(j)$ and hence, $R_j > 2c_{\text{av}}(j)$. Therefore, $c_{\text{av}}(j) = \sum_{i \in \mathcal{F}_j} x_{ij} c_{ij} \geq \sum_{i \in \mathcal{F}_j \setminus \mathcal{U}_j} x_{ij} c_{ij} \geq R_j \cdot \sum_{i \in \mathcal{F}_j \setminus \mathcal{U}_j} x_{ij} \geq R_j \cdot \text{vol}(\mathcal{F}_j \setminus \mathcal{U}_j)$ where the last inequality follows because $x_{ij} = y_i$ for all $i \in \mathcal{F}$ and $j \in \mathcal{C}'$. Therefore $\text{vol}(\mathcal{F}_j \setminus \mathcal{U}_j) < 1/2$ and $\text{vol}(\mathcal{U}_j) > 1/2$.

To prove (ii) consider distinct $j, j' \in \mathcal{C}'$. By the definition of R_j we have $c_{jj'} \geq 2R_j$. Hence, for any facility i in $\mathcal{B}(j, R_j)$ we have $c_{ij} < c_{ij'}$, which implies (ii). \square

C Missing Proofs for Section 4.4

Proof of Lemma 4.5. The idea of the proof is to provide for each client an upper bound on the distance C_j (according to the original distance c) traveled by this client. In what follows c_1, c_2 are constants to be determined later.

The upper bound is paid for by two budgets D_j and X_j . The “deterministic” budget D_j is $c_1 T$. The “probabilistic” budget X_j is a random variable (depending on the random choices made by the algorithm).

We will show below that (by a suitable choice of X_j) the connection cost C_j of j can actually be upper bounded by $D_j + X_j$ and $\mathbb{E}[X_j] \leq c_2 c_{\text{av}}^T(j)$. If this can be shown then this will complete our proof of a constant-factor approximation. To this end note that at most ℓ clients j will pay the budget $D_j = c_1 T$ because at most ℓ distances are actually accounted for in the objective function. Analogously to the case of dedicated clustering, we obtain:

$$\mathbb{E}[\text{ALG}] \leq D_j \cdot \ell + \sum_{j \in \mathcal{C}} \mathbb{E}[X_j] \leq c_1 \ell \cdot T + c_2 \cdot \sum_{j \in \mathcal{C}} c_{\text{av}}^T(j).$$

To show the claim consider an arbitrary client j with connection cost C_j . We incrementally construct our upper bound on C_j starting with 0. Each increment will be either charged to D_j or X_j .

Consider a client j and the cluster center j' it is assigned to (possibly $j = j'$). We have that $c_{jj'} \leq 4c_{\text{av}}(j) \leq 4c_{\text{av}}^T(j) + 4T$ by line 8 of Algorithm 3. We charge $4T$ to D_j and $4c_{\text{av}}^T(j)$ with probability 1 to X_j .

We now describe how to pay for the transport from j' to an open facility. There are two cases to distinguish. Either a facility within a radius of βT is opened or not. (Here, $\beta \geq 2$ is a parameter to be determined later.) If yes, then this cost can be covered by charging an additional amount of βT to D_j . In this case the total cost is upper bounded by $D_j = (\beta + 4)T$ and $\mathbb{E}[X_j] = 4c_{\text{av}}^T(j)$.

If no facility within a radius of βT of j' is opened then observe that for all facilities i with $c_{ij'} \geq \beta T$ we have that $c_{ij'}^T = c_{ij'}$ because of $\beta \geq 1$. We now continue to bound the connection cost for this case. Let j'' be the closest client distinct from j' in \mathcal{C}' . We consider the case where j'' and j' are not matched. (The case where they are matched is simpler.) Let j''' be the client in \mathcal{C}' to which j'' is matched i.e. $(j'', j''') \in \mathcal{M}$. By the dependent rounding process one facility in $\mathcal{U}_{j''} \cup \mathcal{U}_{j'''}$ will be opened. We have that $c_{j''j'} = 2R_{j'} = 2R$ and thus $c_{j''j'''} \leq 2R$ and $R_{j''}, R_{j'''} \leq R$ (otherwise, j'' and j''' would not have been matched). This means that in case no facility is opened in the bundle $\mathcal{U}_{j'}$ the client j travels an additional distance (in expectation) of at most $\max(c_{j''j'''} + R_{j''}, c_{j''j'''} + c_{j''j'''} + R_{j'''}) \leq 2R + 2R + R \leq 5R$.

If a facility is opened in the bundle $\mathcal{U}_{j'}$ then we charge this additional connection cost to X_j . The contribution of the additional connection cost in this case to the expectation of X_j cost is at most

$$\sum_{i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', \beta T)} x_{ij'} c_{ij'} = \sum_{i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', \beta T)} x_{ij'} c_{ij'}^T \leq \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{B}(j', \beta T)} x_{ij'} c_{ij'}^T. \quad (11)$$

Here, equality holds because we assume that no facility is opened in $\mathcal{B}(j', \beta T)$ where $\beta \geq 1$ and because therefore $c_{ij'} = c_{ij'}^T$ for all $i \in \mathcal{U}_{j'} \setminus \mathcal{B}(j', \beta T)$. The right hand side of (11) is denoted by $c_{\text{far}}^T(j')$ and is clearly upper bounded by $\sum_{i \in \mathcal{F}_{j'}} x_{ij} c_{ij}^T = c_{\text{av}}^T(j')$.

We finally handle the case where no facility in $\mathcal{U}_{j'}$ is opened and where j additionally travels a distance of at most $5R$. If $R \leq \beta T$, we can charge the additional travel distance of at most $5\beta T$ to D_j . Hence, we focus on the difficult case where $R > \beta T$ and where the maximum distance traveled can be unbounded in terms of T . We charge this additional cost to X_j . We bound the probability that this case occurs. We claim that $\text{vol}(\mathcal{U}_{j'})$ is at least $1 - c_{\text{far}}^T(j')/R$. To see this, note that for all facilities in $\mathcal{F}_{j'} \setminus \mathcal{U}_{j'}$ we have that $c_{ij'}^T = c_{ij'}$ because $R > \beta T$ and $\beta \geq 1$. Hence

$$\begin{aligned} c_{\text{far}}^T(j') &= \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{B}(j', \beta T)} x_{ij'} \cdot c_{ij'}^T = \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{B}(j', \beta T)} x_{ij'} \cdot c_{ij'} \geq \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{U}_{j'}} x_{ij'} \cdot c_{ij'} \\ &\geq R \cdot \sum_{i \in \mathcal{F}_{j'} \setminus \mathcal{U}_{j'}} x_{ij'} = R \cdot \text{vol}(\mathcal{F}_{j'} \setminus \mathcal{U}_{j'}) = R \cdot (1 - \text{vol}(\mathcal{U}_{j'})), \end{aligned}$$

which implies the claim. Here, note that $\mathcal{B}(j', \beta T) \subseteq \mathcal{U}_{j'}$ because $R > \beta T$. This means that j travels the additional distance of at most $5R$ with probability at most $c_{\text{far}}^T(j')/R$ and hence the increment to X_j in expectation is upper bounded by $5 \cdot c_{\text{far}}^T(j')$. Thus, for the case where no facility is opened within a radius of βT around j' , we can upper bound $\mathbb{E}[X_j]$ by:

- an expected cost of serving client j through the closest cluster center j' that is $4 \cdot c_{\text{av}}^T(j)$ (random part), plus
- a value $c_{\text{far}}^T(j')$ (with probability at most one), plus
- a value $5 \cdot R$ with probability at most $c_{\text{far}}^T(j')/R$.

Hence $\mathbb{E}[X_j] \leq 4 \cdot c_{\text{av}}^T(j) \cdot 1 + c_{\text{far}}^T(j') \cdot 1 + 5 \cdot R \cdot c_{\text{far}}^T(j')/R = 4c_{\text{av}}^T(j) + 5c_{\text{far}}^T(j')$. As in the oblivious clustering we sort the clients according to c_{av} rather than c_{av}^T we do not necessarily have that $c_{\text{av}}^T(j')$ or even $c_{\text{far}}^T(j')$ are upper bounded by $c_{\text{av}}^T(j)$. We still can relate the latter two quantities in the following way.

First, assume that $c_{jj'} > \alpha T$ where $1 \leq \alpha < \beta - 1$ is a parameter to be determined later. We have that $c_{\text{av}}(j') \leq c_{\text{av}}(j)$ by our (oblivious) clustering. Hence $c_{\text{av}}^T(j') \leq c_{\text{av}}(j') \leq c_{\text{av}}(j) \leq c_{\text{av}}^T(j) + T$. On the other hand, $\alpha T < c_{jj'} \leq 4c_{\text{av}}(j)$ since j was assigned to j' . Hence $T < 4/\alpha \cdot c_{\text{av}}(j)$ and thus $c_{\text{av}}^T(j') \leq (1 + 4/\alpha)c_{\text{av}}^T(j)$. Since $c_{\text{far}}^T(j') \leq c_{\text{av}}^T(j')$ we can upper bound $\mathbb{E}[X_j]$ in this case by $(9 + 20/\alpha)c_{\text{av}}^T(j)$.

Second, assume that $c_{jj'} \leq \alpha T$. Recall that we assume further that no facility is opened within $\mathcal{B}(j', \beta T)$. We claim that in the assignment vector x the total demand assigned from j' to $\mathcal{F} \setminus \mathcal{B}(j', \beta T)$ is *at most* the total demand assigned from j to $\mathcal{F} \setminus \mathcal{B}(j', \beta T)$. This is, because any facility within the ball $\mathcal{B}(j', \beta T)$ is (trivially) strictly closer than any facility not in this ball. Hence, if j would manage to assign strictly more demand to facilities inside the ball than j' does, then we could construct a new assignment for j' that also serves strictly more demand of j' within this ball contradicting the optimality of x . Now, we are going to construct a (potentially suboptimal) assignment of the part of the demand of j' contributing to $c_{\text{far}}^T(j')$ that can be upper bounded in terms of $c_{\text{av}}^T(j)$. As the optimum assignment will clearly will have the same upper bound this will conclude our proof. To this end, we simply assign the demand of j' outside of the ball $\mathcal{B}(j', \beta T)$ in the same way as does j . Note that by our above claim this provides enough demand as j ships at least as much demand outside the ball as j' does. In particular let i be an arbitrary facility outside the ball. We now set $x'_{ij'} := x_{ij}$ to obtain our new assignment for j' . Note that by triangle inequality $c_{ij} \geq c_{ij'} - c_{jj'} \geq (\beta - \alpha)T \geq T$ and thus $c_{ij} = c_{ij}^T$ (a constraint $\alpha \leq \beta - 1$ was introduced to obtain $c_{ij} = c_{ij}^T$ in this case). Therefore

$$\frac{c_{ij'}^T}{c_{ij}^T} = \frac{c_{ij'}}{c_{ij}} \leq \frac{c_{ij'}}{c_{ij'} - c_{jj'}} \leq \frac{\beta T}{(\beta - \alpha)T} = \frac{\beta}{\beta - \alpha}.$$

x' can be not optimal assignment for j' , hence

$$c_{\text{far}}^T(j') \leq \sum_{i \in \mathcal{F} \setminus \mathcal{B}(j', \beta T)} x'_{ij'} c_{ij'}^T \leq \frac{\beta}{\beta - \alpha} \sum_{i \in \mathcal{F} \setminus \mathcal{B}(j', \beta T)} x_{ij} c_{ij}^T \leq \frac{\beta}{\beta - \alpha} c_{\text{av}}^T(j).$$

In the end we have two upper bounds for both budgets:

$$D_j \leq (4 + 5\beta)T, \\ \mathbb{E}[X_j] \leq \max \left\{ 4 + \frac{5\beta}{\beta - \alpha}, 9 + \frac{20}{\alpha} \right\} c_{\text{av}}^T(j).$$

Plugging $\alpha = 2$ and $\beta = 3$ gives the desired constants in the claim. \square

D Missing Proofs from Section 5

Corollary D.1. *Let \mathcal{I} will be an instance of ORDERED k -MEDIAN with constant number of different weights in w . There exists a randomized algorithm that solves ORDERED k -MEDIAN on \mathcal{I} and gives 38-approximation (in expectation) in polynomial time.*

Proof. We have $\mathcal{O}(|\{w_j : j \in \{1, 2, \dots, n\}\}|) = \mathcal{O}(1)$. Therefore using Lemmas 5.1 and 4.5 we get 38-approximation solution in $n^{\mathcal{O}(1)} \cdot t(\mathcal{A}) = n^{\mathcal{O}(1)} \cdot \text{poly}(nm) = \text{poly}(nm)$ time. \square

In Lemma D.4 we show how to reduce the number of different weights to at most $\mathcal{O}(\log_{1+\epsilon} n)$. Main idea of such reduction is partitioning an interval $[0, w_1]$ into buckets with geometrical step $1 + \epsilon$. Solving such instance we lose factor $1 + \epsilon$ on approximation because for α -approximation solution W_α^* for \mathcal{I}^* , optimal solution W_{OPT}^* for \mathcal{I}^* and optimal solution W_{OPT} for \mathcal{I} we have

$$\begin{aligned} \text{cost}_{\mathcal{I}}(W_\alpha^*) &\leq (1 + \epsilon) \text{cost}_{\mathcal{I}^*}(W_\alpha^*) \leq (1 + \epsilon) \alpha \cdot \text{cost}_{\mathcal{I}^*}(W_{\text{OPT}}^*) \leq (1 + \epsilon) \alpha \cdot \text{cost}_{\mathcal{I}^*}(W_{\text{OPT}}) \leq \\ &\leq (1 + \epsilon) \alpha \cdot \text{cost}_{\mathcal{I}}(W_{\text{OPT}}). \end{aligned}$$

Theorem D.2. *Let \mathcal{A} be an algorithm for RECTANGULAR ORDERED k -MEDIAN described in Section 4.4 (Algorithm 1 with oblivious clustering, i.e., Algorithm 3). Let $t(\mathcal{A})$ be running time of \mathcal{A} . Then for any $\epsilon > 0$ there exists an $(38 + \epsilon)$ -approximation algorithm for ORDERED k -MEDIAN that runs in $(nm)^{\mathcal{O}(\log_{1+\epsilon} n)} \cdot t(\mathcal{A})$ time.*

Proof. We change a vector of weights w into w^* using Lemma D.4 and use Lemma 5.1. □

Corollary D.3. *There exists a randomized algorithm for ORDERED k -MEDIAN that gives $(38 + \epsilon)$ -approximation (in expectation) in quasi-polynomial time $n^{\mathcal{O}(1/\epsilon \log n)} \cdot \text{poly}(m)$ for any $\epsilon > 0$.*

Proof. Using Theorem D.2 and Lemma 4.5 we get an $38(1 + \epsilon)$ -approximation solution in $n^{\mathcal{O}(\log_{1+\epsilon} n)}$. $\text{poly}(m) = n^{\mathcal{O}(1/\epsilon \log n)} \cdot \text{poly}(m)$ time. □

Lemma D.4. *Let $\mathcal{I} = (\mathcal{F}, \mathcal{C}, c, k, w)$ be an instance of ORDERED k -MEDIAN and $\epsilon > 0$. There exists an instance $\mathcal{I}^* = (\mathcal{F}, \mathcal{C}, c, k, w^*)$ of ORDERED k -MEDIAN such that for any solution $\mathcal{W} \subseteq \mathcal{F}$, $|\mathcal{W}| = k$ we have*

$$\text{cost}_{\mathcal{I}^*}(\mathcal{W}) \leq \text{cost}_{\mathcal{I}}(\mathcal{W}) \leq (1 + \epsilon) \cdot \text{cost}_{\mathcal{I}^*}(\mathcal{W})$$

and w^ has at most $\mathcal{O}(\log_{1+\epsilon} n)$ different values, i.e., $|\{a : \exists j \quad w_j^* = a\}| \in \mathcal{O}(\log_{1+\epsilon} n)$.*

Proof. We define w^* by

$$w_j^* = \begin{cases} w_1 & \text{for } j = 1, \\ (1 + \epsilon)^{\lfloor \log_{1+\epsilon} w_j \rfloor} & \text{for } w_j > \frac{\epsilon w_1}{n} \text{ and } j \neq 1, \\ 0 & \text{for } w_j \leq \frac{\epsilon w_1}{n}. \end{cases} \quad (12)$$

First inequality follows directly from the definition of w_j^* . For the second inequality we have

$$\begin{aligned} \text{cost}_{\mathcal{I}}(\mathcal{W}) &= w_1 \cdot c_1^{\rightarrow}(\mathcal{W}) + \sum_{\substack{j: w_j > \frac{\epsilon w_1}{n} \\ j \neq 1}} w_j \cdot c_j^{\rightarrow}(\mathcal{W}) + \sum_{\substack{j: w_j \leq \frac{\epsilon w_1}{n} \\ j \neq 1}} w_j \cdot c_j^{\rightarrow}(\mathcal{W}) \leq \\ &= w_1^* \cdot c_1^{\rightarrow}(\mathcal{W}) + \sum_{\substack{j: w_j > \frac{\epsilon w_1}{n} \\ j \neq 1}} (1 + \epsilon) \cdot w_j^* \cdot c_j^{\rightarrow}(\mathcal{W}) + \sum_{\substack{j: w_j \leq \frac{\epsilon w_1}{n} \\ j \neq 1}} \frac{\epsilon w_1}{n} \cdot c_j^{\rightarrow}(\mathcal{W}) \leq \\ &= w_1^* \cdot c_1^{\rightarrow}(\mathcal{W}) + \sum_{\substack{j: w_j > \frac{\epsilon w_1}{n} \\ j \neq 1}} (1 + \epsilon) \cdot w_j^* \cdot c_j^{\rightarrow}(\mathcal{W}) + \epsilon \cdot w_1^* \cdot c_1^{\rightarrow}(\mathcal{W}) = \\ &= (1 + \epsilon) \cdot \sum_{j: w_j > \frac{\epsilon w_1}{n}} w_j^* \cdot c_j^{\rightarrow}(\mathcal{W}) = (1 + \epsilon) \cdot \text{cost}_{\mathcal{I}^*}(\mathcal{W}). \end{aligned}$$

Let us assume that there is at least $2 \log_{1+\epsilon} n + 5$ different values w_j^* and n is large enough. We know that the highest value of w_j^* is equal to w_1 . It is possible that the lowest value of w_j^* is

equal to 0. By the induction we can show that the p -th highest value of $\{w_j^* : w_j^* > 0\}$ for $p \in \{2, 3, \dots, \lceil 2\log_{1+\epsilon} n \rceil + 3\}$ is at most $\frac{w_1}{(1+\epsilon)^{p-2}}$. Therefore

$$\min\{w_j^* : w_j^* > 0\} < \frac{w_1}{(1+\epsilon)^{\lceil 2\log_{1+\epsilon} n \rceil + 3 - 2}} \leq \frac{w_1}{(1+\epsilon)n^2} \leq \frac{w_1}{(1+\epsilon)n}.$$

So there exists j such that $\frac{w_1}{n} \geq (1+\epsilon)w_j^* > 0$ and $w_j > \frac{\epsilon w_1}{n}$. From the definition we have $(1+\epsilon)w_j^* = (1+\epsilon)^{\lfloor \log_{1+\epsilon} w_j \rfloor + 1} > (1+\epsilon)^{\log_{1+\epsilon} w_j} = w_j > \frac{\epsilon w_1}{n} > \frac{w_1}{n}$. Contradiction. Therefore it is not possible that w^* has at least $2\log_{1+\epsilon} n + 5$ different values. It means that w^* has at most $\mathcal{O}(\log_{1+\epsilon} n)$ different values. \square

E Details for Polynomial-Time Algorithm from Section 5

We apply the clever bucketing approach of Aouad and Segev [3], which allows guessing an approximate reduced cost function in polynomial time. Applying this idea within our LP-based framework, however, renders the analysis more intricate than the one with quasi-polynomial time.

E.1 Distance Bucketing

Let \mathcal{W}_{OPT} be an optimal solution to a given ORDERED k -MEDIAN instance. Let $c_{\max} := c_1^{\rightarrow}(\mathcal{W}_{\text{OPT}})$ be the maximum connection cost in this solution. We assume that we know c_{\max} as it is one of $\mathcal{O}(mn)$ many possible distances in the input. Fix an error parameter $\epsilon > 0$ and let $c_{\min} := \epsilon \cdot c_{\max}/n$. Roughly speaking, distances smaller than c_{\min} can have only negligible impact on any feasible solution as they may increase its cost by a factor of at most $1 + \epsilon$.

We now partition the distances of the vector $c^{\rightarrow}(\mathcal{W}_{\text{OPT}})$ into $S := \lceil \log_{1+\epsilon}(n/\epsilon) \rceil = \mathcal{O}(\frac{1}{\epsilon} \log \frac{n}{\epsilon})$ many distance classes. More precisely, for all $s = 0, \dots, S-1$ introduce the intervals $D_s = (c_{\max}(1+\epsilon)^{-(s+1)}, c_{\max}(1+\epsilon)^{-s}]$. Let $D_S = [0, c_{\max}(1+\epsilon)^{-S}] \ni c_{\min}$. For all $s = 0, \dots, S$ let $J_s = \{j \mid c_j^{\rightarrow}(\mathcal{W}_{\text{OPT}}) \in D_s\}$ and let $C_s = \{c_j^{\rightarrow}(\mathcal{W}_{\text{OPT}}) \mid j \in J_s\}$. The classes C_0, \dots, C_S form a disjoint partition of $c^{\rightarrow}(\mathcal{W}_{\text{OPT}})$ where some of the classes may, however, be empty. For technical reasons, we assume that none of the input distances $c_{ij}, i \in \mathcal{F}, j \in \mathcal{C}$ coincides with a boundary of one of the intervals D_s for some $s = 0, \dots, S$. This can be achieved by slightly increasing all boundaries of the intervals using the fact that the intervals are left-open. Additionally we define $J_{\geq s} = \bigcup_{r=s}^S J_r$.

E.2 Guessing Average Weights

For any non-empty class C_s let

$$w_{\text{av}}^s := \frac{1}{|C_s|} \sum_{j \in J_s} w_j \tag{13}$$

denote the *average* weight applied to distances in this class. If C_s is empty then w_{av}^s denotes the smallest weight w_j applied to some distance $c_j^{\rightarrow}(\mathcal{W}_{\text{OPT}})$ in a non-empty class C_l with $l < s$. Such a class always exists as $C_0 \ni c_{\max}$ is non-empty.

As argued by Aouad and Segev [3], it is possible to guess the values of w_{av}^s up to a factor of $1 + \epsilon$ in polynomial time $n^{\mathcal{O}(1/\epsilon \log 1/\epsilon)}$. This is, because we have $w_{\text{av}}^0 \geq w_{\text{av}}^1 \geq \dots \geq w_{\text{av}}^S$ and because it suffices to guess those values as powers of $1 + \epsilon$. More precisely, as a result of this we assume that we are given values $w_{\text{gs}}^0 \geq w_{\text{gs}}^1 \geq \dots \geq w_{\text{gs}}^S$ with $w_{\text{av}}^s \leq w_{\text{gs}}^s \leq (1+\epsilon)w_{\text{av}}^s$ for $i = 0, \dots, S$.

E.3 Reduced Cost Function and LP-Solving

We are now ready to define our reduced cost function. For all values of $d \in [0, c_{\max}]$ let $w(d)$ be the weight w_{gs}^s such that $d \in D_s$ for some $s \in \{0, \dots, S\}$. For each $i \in \mathcal{F}, j \in \mathcal{C}$ such that $c_{ij} \leq c_{\max}$ let $c_{ij}^r := w(c_{ij}) \cdot c_{ij}$. Now solve the linear program $\text{LP}(c^r)$ with additional constraints $x_{ij} = 0$ for all $i \in \mathcal{F}, j \in \mathcal{C}$ such that $c_{ij} > c_{\max}$. In what follows let (x, y) denote an optimal solution to this LP. Now apply the rounding algorithm of Charikar and Li with oblivious clustering (Algorithm 1 with clustering as in Algorithm 3) to obtain an integral solution $A \subseteq \mathcal{F}, |A| = k$.

Let OPT be the value (cost) of an optimum solution W_{OPT} for ORDERED k -MEDIAN and let OPT^* be the value of an optimum solution for $\text{LP}(c^r)$, let A be the solution for ORDERED k -MEDIAN computed by our algorithm. We define distance class (interval) in which the distance d falls by $D(d)$ and $w_{n+1} = 0$.

Using Lemma 4.5 with $T_\ell = \max(D(c_\ell^\rightarrow(W_{\text{OPT}})))$ for each $\ell = 1, \dots, n$ we obtain

$$\mathbb{E}[\text{cost}_\ell(A)] \leq c_1 \cdot \ell T_\ell + c_2 \cdot \sum_{j \in \mathcal{C}} c_{\text{av}}^{T_\ell}(j). \quad (14)$$

We can partition the cost $\text{cost}(A)$ of our algorithm as follows into rectangular pieces

$$\begin{aligned} \mathbb{E}[\text{cost}(A)] &= \mathbb{E} \left[\sum_{\ell=1}^n w_\ell \cdot c_\ell^\rightarrow(A) \right] = \mathbb{E} \left[\sum_{\ell=1}^n \sum_{r=1}^\ell (w_\ell - w_{\ell+1}) \cdot c_r^\rightarrow(A) \right] = \\ &= \mathbb{E} \left[\sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot \text{cost}_\ell(A) \right] \stackrel{(14)}{\leq} 19 \cdot \sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot \ell T_\ell + 19 \cdot \sum_{\ell=1}^n \sum_{j \in \mathcal{C}} (w_\ell - w_{\ell+1}) \cdot c_{\text{av}}^{T_\ell}(j). \end{aligned} \quad (15)$$

We would like to upper bound this in terms of OPT . We know that the optimal solution pays at least cost $\inf(D_s)$ for each distance in distance bucket C_s and thus

$$\text{OPT} \geq \sum_{s=0}^S \left(\inf(D_s) \cdot \sum_{\ell \in J_s} w_\ell \right) = \sum_{s=0}^S \inf(D_s) \cdot w_{\text{av}}^s \cdot |C_s|. \quad (16)$$

Lemma E.1.

$$\sum_{s=0}^S \inf(D_s) \cdot w_{\text{av}}^s \cdot |C_s| \geq \frac{1}{1+\epsilon} \sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot \ell \cdot T_\ell. \quad (17)$$

Proof. The right hand side is equal to

$$\begin{aligned} &\frac{1}{1+\epsilon} \sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot \ell \cdot \max(D(c_\ell^\rightarrow(W_{\text{OPT}}))) \leq \\ &\sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot \ell \cdot \inf(D(c_\ell^\rightarrow(W_{\text{OPT}}))) = \sum_{s=0}^S \sum_{\ell \in J_s} (w_\ell - w_{\ell+1}) \cdot \ell \cdot \inf(D_s) = \\ &\sum_{\substack{s=0 \\ J_s \neq \emptyset}}^S \inf(D_s) \left(w_{\min(J_s)} \cdot \min(J_s) + \sum_{\ell \in J_s \setminus \min(J_s)} w_\ell \cdot \ell - \sum_{\ell \in J_s \setminus \max(J_s)} w_{\ell+1} \cdot \ell - w_{\max(J_s)+1} \cdot \max(J_s) \right) = \\ &\sum_{\substack{s=0 \\ J_s \neq \emptyset}}^S \inf(D_s) \left(w_{\min(J_s)} \cdot \min(J_s) + \sum_{\ell \in J_s \setminus \min(J_s)} w_\ell \cdot \ell - \sum_{\ell \in J_s \setminus \min(J_s)} w_\ell \cdot (\ell - 1) - w_{\max(J_s)+1} \cdot \max(J_s) \right) = \end{aligned}$$

$$\begin{aligned}
& \sum_{\substack{s=0 \\ J_s \neq \emptyset}}^S \inf(D_s) \left(w_{\min(J_s)} \cdot (\min(J_s) - 1) + \sum_{\ell \in J_s} w_\ell - w_{\max(J_s)+1} \cdot \max(J_s) \right) = \\
& \sum_{s=0}^S \inf(D_s) \cdot w_{\text{av}}^s \cdot |C_s| + \sum_{\substack{s=0 \\ J_s \neq \emptyset}}^S \inf(D_s) (w_{\min(J_s)} \cdot (\min(J_s) - 1) - w_{\max(J_s)+1} \cdot \max(J_s)).
\end{aligned}$$

Proof ends with showing that the second factor is non-positive:

$$\begin{aligned}
& \sum_{\substack{s=0 \\ J_s \neq \emptyset}}^S \inf(D_s) (w_{\min(J_s)} \cdot (\min(J_s) - 1) - w_{\max(J_s)+1} \cdot \max(J_s)) = \\
& \sum_{\substack{s=1 \\ J_s \neq \emptyset}}^S \inf(D_s) \cdot w_{\min(J_s)} \cdot (\min(J_s) - 1) - \sum_{\substack{s=0 \\ J_s \neq \emptyset}}^{S-1} \inf(D_s) \cdot w_{\max(J_s)+1} \cdot \max(J_s) = \\
& \sum_{\substack{s=1 \\ J_s \neq \emptyset}}^S \inf(D_s) \cdot w_{\min(J_s)} \cdot (\min(J_s) - 1) - \sum_{\substack{s=1 \\ J_{s-1} \neq \emptyset}}^S \inf(D_{s-1}) \cdot w_{\max(J_{s-1})+1} \cdot \max(J_{s-1}) = \\
& \sum_{\substack{s=1 \\ J_s \neq \emptyset}}^S \inf(D_s) \cdot w_{\min(J_{\geq s})} \cdot (\min(J_{\geq s}) - 1) - \sum_{\substack{s=1 \\ J_{s-1} \neq \emptyset}}^S \inf(D_{s-1}) \cdot w_{\min(J_{\geq s})} \cdot (\min(J_{\geq s}) - 1) = \quad (18) \\
& \sum_{\substack{s=1 \\ J_{s-1} \neq \emptyset \\ J_s \neq \emptyset}}^S (\inf(D_s) - \inf(D_{s-1})) \cdot w_{\min(J_{\geq s})} \cdot (\min(J_{\geq s}) - 1) + \\
& \sum_{\substack{0 \leq s_1 < s_2 \leq S \\ s_1+1 < s_2 \\ J_{s_1}, J_{s_2} \neq \emptyset \\ J_{s_3} = \emptyset \text{ for } s_1 < s_3 < s_2}} \inf(D_{s_2}) \cdot w_{\min(J_{\geq s_2})} \cdot (\min(J_{\geq s_2}) - 1) - \inf(D_{s_1}) \cdot w_{\min(J_{\geq s_1+1})} \cdot (\min(J_{\geq s_1+1}) - 1) \leq \\
& \sum_{\substack{0 \leq s_1 < s_2 \leq S \\ s_1+1 < s_2 \\ J_{s_1}, J_{s_2} \neq \emptyset \\ J_{s_3} = \emptyset \text{ for } s_1 < s_3 < s_2}} \inf(D_{s_2}) \cdot w_{\min(J_{\geq s_2})} \cdot (\min(J_{\geq s_2}) - 1) - \inf(D_{s_1}) \cdot w_{\min(J_{\geq s_2})} \cdot (\min(J_{\geq s_2}) - 1) = \\
& \sum_{\substack{0 \leq s_1 < s_2 \leq S \\ s_1+1 < s_2 \\ J_{s_1}, J_{s_2} \neq \emptyset \\ J_{s_3} = \emptyset \text{ for } s_1 < s_3 < s_2}} (\inf(D_{s_2}) - \inf(D_{s_1})) \cdot w_{\min(J_{\geq s_2})} \cdot (\min(J_{\geq s_2}) - 1) \leq 0.
\end{aligned}$$

The equality (18) is just a merge of two sums into two cases: when two consecutive class C_{s-1}, C_s are non-empty or there are some positive number of empty classes between two non-empty classes C_{s_1}, C_{s_2} . \square

For the second term from (15) we have

$$\sum_{\ell=1}^n \sum_{j \in C} (w_\ell - w_{\ell+1}) \cdot c_{\text{av}}^{T_\ell}(j) =$$

$$\begin{aligned}
& \sum_{\ell=1}^n \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} (w_\ell - w_{\ell+1}) \cdot x_{ij} \cdot c_{ij}^{T_\ell} = \\
& \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot \sum_{\ell=1}^n (w_\ell - w_{\ell+1}) \cdot c_{ij}^{T_\ell} \stackrel{(6)}{=} \\
& \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot \sum_{\substack{\ell=1 \\ \ell: c_{ij} > T_\ell}}^n (w_\ell - w_{\ell+1}) \cdot c_{ij} = \\
& \sum_{s=0}^S \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij} \cdot c_{ij} \cdot \sum_{\substack{\ell=1 \\ \ell \in J_{\geq s+1}}}^n (w_\ell - w_{\ell+1}) = \\
& \sum_{s=0}^S \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij} \cdot c_{ij} \cdot w_{\min\{J_{\geq s+1}\}} \leq \\
& \sum_{s=0}^S \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij} \cdot c_{ij} \cdot w_{\text{av}}^s \leq (1 + \epsilon) \sum_{s=0}^S \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij} \cdot c_{ij} \cdot w_{\text{gs}}^s = \\
& (1 + \epsilon) \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij} \cdot c_{ij}^r \stackrel{(1)}{=} (1 + \epsilon) \cdot \text{OPT}^*. \tag{19}
\end{aligned}$$

This we can upper bound in terms of value of optimal solution OPT. For that let us define the optimal solution W_{OPT} as a feasible solution of $LP(c^r)$ and denote it as $(x^{\text{OPT}}, y^{\text{OPT}})$. It means that $y_i^{\text{OPT}} = 1 \iff i \in W_{\text{OPT}}$ and $y_i^{\text{OPT}} = 0 \iff i \notin W_{\text{OPT}}$.

$$\begin{aligned}
\text{OPT}^* & \leq \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij}^{\text{OPT}} c_{ij}^r = \sum_{j \in \mathcal{C}} \sum_{i \in \mathcal{F}} x_{ij}^{\text{OPT}} c_{ij} w(c_{ij}) = \sum_{s=0}^S \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij}^{\text{OPT}} c_{ij} w_{\text{gs}}^s \leq \\
& \sum_{s=0}^S \left(\max(D_s) \cdot w_{\text{gs}}^s \sum_{\substack{j \in \mathcal{C}, i \in \mathcal{F} \\ c_{ij} \in C_s}} x_{ij}^{\text{OPT}} \right) = \\
& \sum_{s=0}^S \max(D_s) \cdot w_{\text{gs}}^s \cdot |C_s| \leq (1 + \epsilon)^2 \cdot \sum_{s=0}^S \inf(D_s) \cdot w_{\text{av}}^s \cdot |C_s| = \\
& (1 + \epsilon)^2 \cdot \sum_{s=0}^S \sum_{\ell \in J_s} w_\ell \cdot \inf(D_s) \leq (1 + \epsilon)^2 \cdot \text{OPT}. \tag{20}
\end{aligned}$$

In the end we have

$$\text{cost}(A) \stackrel{(15),(16),(17),(19),(20)}{\leq} (1 + \epsilon)^3 \cdot 38 \cdot \text{OPT}. \tag{21}$$