

Facility Location on High-Dimensional Euclidean Spaces

Euiwoong Lee ✉

University of Michigan, Ann Arbor, MI, USA

Kijun Shin ✉

Seoul National University, South Korea

Abstract

Recent years have seen great progress in the approximability of fundamental clustering and facility location problems on high-dimensional Euclidean spaces, including k -Means and k -Median. While they admit strictly better approximation ratios than their general metric versions, their approximation ratios are still higher than the hardness ratios for general metrics, leaving the possibility that the ultimate optimal approximation ratios will be the same between Euclidean and general metrics. Moreover, such an improved algorithm for Euclidean spaces is not known for Uncapacitated Facility Location (UFL), another fundamental problem in the area.

In this paper, we prove that for any $\gamma \geq 1.6774$ there exists $\varepsilon > 0$ such that Euclidean UFL admits a $(\gamma, 1 + 2e^{-\gamma} - \varepsilon)$ -bifactor approximation algorithm, improving the result of Byrka and Aardal [3]. Together with the $(\gamma, 1 + 2e^{-\gamma})$ NP-hardness in general metrics, it shows the first separation between general and Euclidean metrics for the aforementioned basic problems. We also present an $(\alpha_{\text{Li}} - \varepsilon)$ -(unifactor) approximation algorithm for UFL for some $\varepsilon > 0$ in Euclidean spaces, where $\alpha_{\text{Li}} \approx 1.488$ is the best-known approximation ratio for UFL by Li [15].

2012 ACM Subject Classification Theory of computation \rightarrow Facility location and clustering

Keywords and phrases Approximation Algorithms, Clustering, Facility Location

Digital Object Identifier 10.4230/LIPIcs.ITCS.2025.70

Related Version *Full Version*: <https://arxiv.org/abs/2501.18105> [14]

Funding *Euiwoong Lee*: Supported by organization NSF CCF-2236669 and a gift from Google.

1 Introduction

The (metric) Uncapacitated Facility Location (UFL) is one of the most fundamental problems in computer science and operations research. The input of the problem consists of a metric space (X, d) , a set of facility locations $\mathcal{F} \subseteq X$, a set of clients $\mathcal{C} \subseteq X$, as well as facility opening costs $\{f_i\}_{i \in \mathcal{F}}$. The goal is open a subset of centers $S \subseteq \mathcal{F}$ to minimize the sum of the *opening cost* ($\sum_{i \in S} f_i$) and the *connection cost* ($\sum_{j \in \mathcal{C}} \min_{i \in S} d(i, j)$). After intensive research efforts over the years [11, 12, 16, 15], the best approximation ratio is 1.488 [15] and the best hardness ratio is 1.463 [11].

As the objective function is the sum of two heterogeneous terms of the opening cost and the connection cost, the natural notion of *bifactor approximation* has been actively studied as well. Formally, given an instance of UFL, a solution $S \subseteq \mathcal{F}$ is called an (λ_f, λ_c) -approximation for some $\lambda_f, \lambda_c \geq 1$ if, for any solution $T \subseteq \mathcal{F}$, the total cost of S is at most $\lambda_f \cdot F^* + \lambda_c \cdot C^*$, where F^*, C^* denote the opening and connection cost of T respectively. In particular, the case $\lambda_f = 1$, also known as a λ_c -*Lagrangian Multiplier Preserving (LMP)* approximation, has been actively studied due to its connection to another fundamental clustering problem of k -Median. There is a $(2 - \varepsilon)$ -LMP approximation for some $\varepsilon > 2 \cdot 10^{-7}$ [8], and any λ_c -LMP approximation for UFL can be translated to $1.307 \cdot \lambda_c$ -approximation for k -Median [9].



© Euiwoong Lee and Kijun Shin;

licensed under Creative Commons License CC-BY 4.0

16th Innovations in Theoretical Computer Science Conference (ITCS 2025).

Editor: Raghu Meka; Article No. 70; pp. 70:1–70:21

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Generalizing the hardness of Guha and Khuller [11], Jain, Mahdian and Saberi [12] proved that no (λ_f, λ_c) -approximation polynomial-time algorithm exists for $\lambda_c < 1 + 2e^{-\lambda_f}$ unless $\mathbf{P} = \mathbf{NP}$. (Guha-Khuller's hardness ratio $\gamma \approx 1.463$ is exactly the solution of $\gamma = 1 + 2e^{-\gamma}$.) While the optimal value for λ_c is not known for small values of λ_f , Byrka and Aardal gave an algorithm that achieves an $(\lambda_f, 1 + 2e^{-\lambda_f})$ -approximation for any $\lambda_f \geq 1.6774$ [3].

Euclidean spaces are arguably the most natural metric spaces for facility location and clustering problems. Formally, *Euclidean UFL* is a special case of UFL where the underlying metric is $(\mathbb{R}^k, \|\cdot\|_2)$ for some dimension k . When $k = O(1)$, this problem admits a PTAS [5], while the problem remains APX-hard when k is part of the input [7].¹

Recent years have seen active studies on related k -Means and k -Median on high-dimensional Euclidean spaces [1, 10, 4], so that the best-known approximation ratios for them are 5.912 and 2.406 respectively. While they are strictly lower than the best-known approximation ratios for general metric spaces (which are 9 and 2.613), they are still larger than the best-known hardness ratios for general metrics (which are $1 + 8/e \approx 3.943$ and $1 + 2/e \approx 1.73$) [12], which means that it is still plausible that the optimal approximation ratios for k -Median and k -Means are the same between Euclidean metrics and general metrics.

Our first result is the first strict separation between Euclidean and general metric spaces for UFL. In particular, we show that Euclidean UFL admits a $(1.6774, 1 + 2e^{-1.6774} - \varepsilon)$ -approximation for some universal constant $\varepsilon > 0$, which is NP-hard to do in general metrics.

► **Theorem 1.** *There exists a $(1.6774, 1 + 2e^{-1.6774} - \varepsilon)$ -approximation algorithm for Euclidean UFL for some $\varepsilon \geq 3 \cdot 10^{-42}$.*

By the result of Mahdian et al. [16], it implies an $(\gamma, 1 + 2e^{-\gamma} - \varepsilon e^{-(\gamma-1.6774)})$ -approximation for any $\gamma \geq 1.6774$. Using this result, we are able to slightly improve the approximation ratio for the best-known $(\alpha_{\text{Li}} \approx 1.488)$ -unifactor approximation of Li [15].

► **Theorem 2.** *There exists a $(\alpha_{\text{Li}} - \varepsilon)$ -approximation algorithm for Euclidean UFL for some $\varepsilon \geq 2 \cdot 10^{-45}$.*

Recent years also have seen great progress on hardness of approximation for clustering problems in high-dimensional Euclidean spaces, including Euclidean k -Means and k -Median [6, 7]. We show that similar techniques extend to UFL as well, proving the APX-hardness.

► **Theorem 3.** *Euclidean UFL is APX-hard.*

2 High-level Plan

Our work is based on the framework of Byrka and Aardal [3] who achieved an optimal $(\lambda_f, 1 + 2e^{-\lambda_f})$ -bifactor approximation for $\lambda_f \geq 1.6774$ in general metrics. We first review their framework. It is based on the following standard linear programming (LP) relaxation:

$$\begin{array}{ll}
 \text{Minimize} & \sum_{i \in \mathcal{F}, j \in \mathcal{C}} d(i, j)x_{ij} + \sum_{i \in \mathcal{F}} f_i y_i \\
 \text{subject to} & \sum_{i \in \mathcal{F}} x_{ij} = 1 \qquad \qquad \qquad \forall j \in \mathcal{C}, \\
 & x_{ij} \leq y_i \qquad \qquad \qquad \forall i \in \mathcal{F}, j \in \mathcal{C}, \\
 & x_{ij}, y_i \in [0, 1] \qquad \qquad \forall i \in \mathcal{F}, j \in \mathcal{C}.
 \end{array}$$

¹ While the cited paper only studies k -Median and k -Means, the soundness analysis in their Theorem 4.1 (of the arXiv version) can be directly extended to any number of open facilities k , implying APX-hardness of Euclidean UFL.

The dual formulation is as follows:

$$\begin{aligned}
& \text{Maximize} && \sum_{j \in \mathcal{C}} v_j \\
& \text{subject to} && \sum_{j \in \mathcal{C}} w_{ij} \leq f_i && \forall i \in \mathcal{F}, \\
& && v_j \leq w_{ij} && \forall i \in \mathcal{F}, j \in \mathcal{C}, \\
& && w_{ij} \geq 0 && \forall i \in \mathcal{F}, j \in \mathcal{C}.
\end{aligned}$$

A feasible solution (x, y) induces the *support graph*, which is defined as the bipartite graph $G = ((\mathcal{F}, \mathcal{C}), E)$ where nodes $i \in \mathcal{F}$ and $j \in \mathcal{C}$ are adjacent iff the corresponding LP variable $x_{ij} > 0$. Two clients $j, j' \in \mathcal{C}$ are considered *neighbors* in G if they share the same facility.

Let (x^*, y^*) be a fixed optimal solution to the primal program. The overall cost is divided into the facility cost $F^* = \sum_{i \in \mathcal{F}} f_i y_i^*$ and the connection cost $C^* = \sum_{i \in \mathcal{F}, j \in \mathcal{C}} d(i, j) x_{ij}^*$. Our goal is to round this solution to obtain a solution S whose total cost is at most $\lambda_f F^* + \lambda_c C^*$; it is well known that it implies the (λ_f, λ_c) -approximation defined in the introduction by scaling [3], so let us redefine the (λ_f, λ_c) -approximation for the rest of the paper so that S is (λ_f, λ_c) -approximate if its total cost is $\lambda_f F^* + \lambda_c C^*$.

The opening cost and connection cost for individual clients can be further divided using the optimal LP dual solution (v^*, w^*) . For each client $j \in \mathcal{C}$, the fractional connection cost is given by $C_j^* = \sum_{i \in \mathcal{F}} d(i, j) x_{ij}^*$, and the fractional facility cost is computed as $F_j^* = v_j^* - C_j^*$. The *irregularity* of the facilities surrounding $j \in \mathcal{C}$ is defined by

$$r_\gamma(j) = \frac{d(j, \mathcal{D}_j) - d(j, \mathcal{D}_j \cup \mathcal{C}_j)}{F_j^*}.$$

Similarly,

$$r'_\gamma(j) = (\gamma - 1) \cdot r_\gamma(j) = \frac{d(j, \mathcal{D}_j \cup \mathcal{C}_j) - d(j, \mathcal{C}_j)}{F_j^*}.$$

If $\sum_{i \in \mathcal{F}'} y_i^* = 0$, we set $d(j, \mathcal{F}') = 0$. Similarly, when $F_j^* = 0$, we define $r_\gamma(j) = 0$ and $r'_\gamma(j) = 0$. According to the definition, the following conditions hold: the irregularity $0 \leq r_\gamma(j) \leq 1$, the average distance to a close facility $C_j = d(j, \mathcal{C}_j) = C_j^* - r'_\gamma(j) \cdot F_j^*$, and the average distance to a distant facility $D_j = d(j, \mathcal{D}_j) = C_j^* + r_\gamma(j) \cdot F_j^*$. The maximum distance to a close facility is bounded by $M_j \leq D_j$.

Clustering of [3]. At a high level, clustering operates based on the support graph $G = ((\mathcal{F}, \mathcal{C}), E)$. For each $c \in \mathcal{C}$, let $N_c := \{c' \in \mathcal{C} : \exists f \in \mathcal{F} \text{ such that } (c, f), (c', f) \in E\}$ be the neighbor of c . The clustering algorithm iteratively selects some client c as a *cluster center*, put all its neighbors into the cluster, and proceed with the remaining clients. Eventually, all clients are partitioned into one of these clusters. After this, for each cluster, exactly one facility adjacent to cluster center is opened. This ensures that every client is connected to a facility that is not too far from them. Therefore, the criteria for choosing cluster centers and opening facilities will determine the quality of solution.

Starting with a fractional solution of the LP (x^*, y^*) and a parameter $\gamma \in (1, 2)$, [3] constructed the *facility-augmented solution* (\bar{x}, \bar{y}) , where each y_i^* value is multiplied by γ and each client $j \in \mathcal{C}$ reconfigures its x_{ij}^* values to be fractionally connected to as close facilities as possible. (E.g., $\bar{x}_{ij} > 0$ implies $x_{ij}^* > 0$, but not vice versa.) With some postprocessing,

70:4 Facility Location on High-Dimensional Euclidean Spaces

one can also assume that $\bar{x}_{ij} \in \{0, \bar{y}_i\}$ for every $i \in \mathcal{F}, j \in \mathcal{C}$. Then one can categorize every facility near $j \in \mathcal{C}$ into two types: *close* facilities $\mathcal{C}_j = \{i \in \mathcal{F} \mid \bar{x}_{ij} > 0\}$ and *distant* facilities $\mathcal{D}_j = \{i \in \mathcal{F} \mid \bar{x}_{ij} = 0 \text{ and } x_{ij}^* > 0\}$. This implies that as γ increases, the clusters become smaller, and more facilities are opened.

Let the *average distance* from $j \in \mathcal{C}$ to a set of facilities $\mathcal{F}' \subseteq \mathcal{F}$ be defined as $d(j, \mathcal{F}') = \left(\sum_{i \in \mathcal{F}'} d(i, j) y_i^* \right) / \left(\sum_{i \in \mathcal{F}'} y_i^* \right)$. Then Let $C_j := d(j, \mathcal{C}_j)$, $M_j := \max_{i \in \mathcal{C}_j} d(j, i)$, and $D_j := d(j, \mathcal{D}_j)$. We have $C_j \leq M_j \leq D_j$.

At this point, the support graph is defined by (\bar{x}, \bar{y}) solution. Intuitively, we choose the client j with the smallest $C_j + M_j$ as a new cluster center. Given this clustering, the standard randomized rounding procedure is as follows:

1. For each cluster center j , choose exactly one facility from its neighboring facility set $\{i \in \mathcal{F} : (i, j) \in E\}$ according to the \bar{y} values. (Recall that the sum of these values is exactly 1.)
2. For any facility $i \in \mathcal{F}$ that is not adjacent to any cluster center in G , independently open i with probability \bar{y}_i .

■ **Algorithm 1** GREEDY: [3]'s clustering algorithm.

Require: Support graph $G = ((\mathcal{F}, \mathcal{C}), E)$

```

L = {}
while C ≠ ∅ do
  c ← arg minj ∈ C (Cj + Mj)
  L ← L ∪ (c, Nc)
  C ← C \ Nc
end while
return L

```

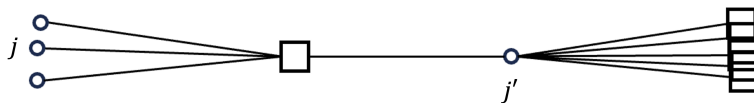
Let us consider one client $j \in \mathcal{C}$ and see how its expected connection cost can be bounded under the above randomized rounding. Byrka and Aardal [3] proved the following properties.

- The probability that at least one facility in \mathcal{C}_j is opened is at least $1 - e^{-1}$.
- The probability that at least one facility in $\mathcal{C}_j \cup \mathcal{D}_j$ is opened is at least $1 - e^{-\gamma}$.
- Let client $j' \in \mathcal{C}$ be a neighbor of j in G . Then, either $\mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j) = \emptyset$ or the *rerouting cost* $d(j, \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)) \leq d(j', j) + d(j', \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)) \leq D_j + C_{j'} + M_{j'}$, holds. Especially, when j' is the cluster center of j , it is at most $C_j + M_j + D_j$. (Li [15] refined this bound to $C_j + (3 - \gamma)M_j + (\gamma - 1)D_j$.)

Then, one can (at least informally) expect that the expected connection cost of j is at most $(1 - e^{-1})C_j + (e^{-1} - e^{-\gamma})D_j + e^{-\gamma}(D_j + C_j + M_j)$. It turns out that setting $\gamma \approx 1.6774$ (the solution of $e^{-1} + e^{-\gamma} - (\gamma - 1)(1 - e^{-1} + e^{-\gamma}) = 0$) ensures that this value is at most $(1 + 2e^{-\gamma})C_j^*$, proving their $(\gamma, 1 + 2e^{-\gamma})$ -bifactor. (See Section 6 for the formal treatment of their analysis as well as our improvement.)

Exploit the Geometry of Euclidean Spaces. In order to strictly improve the approximation ratio, it is natural to attempt to find a cluster N and its center j' where the above inequality holds with some additional slack. Let $cost_{j'}(j) = d(j, \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j))$. Intuitively, our goal is to find a cluster $N \subseteq \mathcal{C}$ with center j' such that

$$\sum_{j \in N} cost_{j'}(j) \leq \sum_{j \in N} ((1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j). \quad (1)$$



■ **Figure 1** A simple case when $\text{cost}_{j'}(j) \approx 3$. There is a client-dense region on the left.

The only requirement from the rounding algorithm is that $N_{j'} \subseteq N$. Compared to [3]’s clustering, we want to shave $\varepsilon_1 \cdot C_j$ on average.

Let us consider the very special case where $C_j = M_j = D_j = 1$ for every $j \in \mathcal{C}$; every facility serving j in the original LP solution (x^*, y^*) is at the same distance from j . Let $j' \in \mathcal{C}$ be a cluster center and $j \in \mathcal{C}$ be in the cluster of j' . Then, a simple *3-hop triangle inequality* (just using $C_j \cap C_{j'} \neq \emptyset$) ensures that $\text{cost}_{j'}(j) \leq 3$, and our goal is to improve it to $(3 - \varepsilon_1)$. If $\text{cost}_{j'}(j) > 3 - \varepsilon_1$, how should the instance look like around j' ?

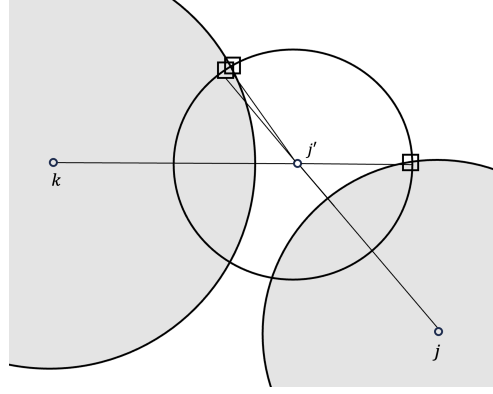
It turns out that the instance around j' must exhibit a very specific structure in order to ensure that the 3-hop triangle inequality is tight for almost every neighbor $j \in N_{j'}$. We must have almost every $j \in N_{j'}$ located around almost the same point at distance 2 from j' , where almost all facility neighbors of j' are at the opposite end of the line connecting $N_{j'}$ and j' . See Figure 1 for an example. Intuitively, the existence of such a *dense region* of clients suggests that if we let a client j'' in the region as a new center, many of the 3-hop-triangle inequalities cannot be tight, which implies average rerouting cost $\text{cost}_{j''}(j') \leq 3 - \varepsilon_1$. If j'' is again problematic, we can repeat this procedure over and over.

However, if we relax the condition to $C_j = M_j = 1 \leq D_j$, certain exceptions begin to emerge. One possible scenario is as follows: Since $\text{cost}_{j'}(j) = d(j, C_{j'} \setminus (C_j \cup D_j))$ captures the rerouting of j to j' ’s close facilities $C_{j'}$, *except the j ’s facilities $C_j \cup D_j$* , if $C_j \cup D_j$ is large enough to exclude the facilities of $C_{j'}$, then $\text{cost}_{j'}(j)$ might not behave as expected. However, a large volume of $C_j \cup D_j$ implies low C_j^*/F_j^* ratio. If the C^*/F^* ratio is sufficiently low, then this *facility-dominant* instance is actually easier to handle with a completely different algorithm, the JMS algorithm [12], which is known to be $(1.11, 1.7764)$ -approximation algorithm.

Therefore, from now on, assume that the cluster centered at j' is connection-dominant. More strictly, assume that for any neighbor j of cluster center j' satisfies that $C_j \cup D_j$ cannot cover half of the ball centered at j' with a radius of 1. At this point, we can finally assert that it is impossible to avoid the formation of a dense region of clients.

Assume towards contradiction that there is no dense region and consider j in j' ’s cluster such that $\text{cost}_{j'}(j) > 3 - \varepsilon_1$. Almost all facilities of j' must be placed in one of two locations: on the opposite side of j relative to j' , or within $C_j \cup D_j$. Since there is no dense region, there must be a neighbor k of j' such that $\text{cost}_{j'}(k) > 3 - \varepsilon_1$, and k is located in a different direction from j . This implies that the facilities positioned on the opposite side of j now help reduce $\text{cost}_{j'}(k)$, forcing that they are in $C_k \cup D_k$; since we assumed that $C_k \cup D_k$ cannot cover half of the unit ball around j' , it implies the angle $\angle jj'k$ must be strictly greater than $\frac{\pi}{2}$! Ultimately, this process can be simplified to the following situation: inserting unit vectors into a unit sphere with every pairwise angle greater than $\frac{\pi}{2} + \varepsilon$ for some constant $\varepsilon > 0$. It is well known that in the geometry of Euclidean space, there is an upper bound $f(\varepsilon)$ on the number of such vectors, and such an upper bound shows that one of the regions around j (or k) we considered must have been dense. See Figure 2 for an example.

However, there are several technical barriers to extending this notion to the general case without restrictions on C_j , M_j , and D_j . In Section 3, we introduce these barriers and formalize the above concept. We also propose sufficient conditions to satisfy (1). From



■ **Figure 2** $C_k \cup D_k$ contains facilities positioned on the opposite side of j .

Section 3.1 to Section 5, we demonstrate how to remove these conditions, leaving only the connection-dominant instance assumptions. In Section 6, we propose and analyze the full algorithm, which achieves improved bi-factor approximation performance.

3 Finding Good Center via Geometry

In this section, we exploit the geometry of Euclidean spaces to prove the existence of a cluster center strictly better than the greedy choice of [3] under certain conditions (Theorem 10). We first define several concepts and introduce their motivation, including the sketch of our algorithm.

Recall that our goal is to find a cluster that satisfies (1). In all the following propositions, γ is a fixed value in the range $\gamma \in (1.6, 2)$.

► **Definition 4.** Suppose $j' \in \mathcal{C}$ is a cluster center. Let $N_{j'}$ be the set of neighbors of j' , and $\varepsilon_1 = 10^{-12}$. Additionally, define two more sets:

$$N_{j'}^- = \{j \in N_{j'} \mid \text{cost}_{j'}(j) > (1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j\}$$

$$N_{j'}^+ = \{j \in \mathcal{C} \mid \text{cost}_{j'}(j) \leq (1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j\}$$

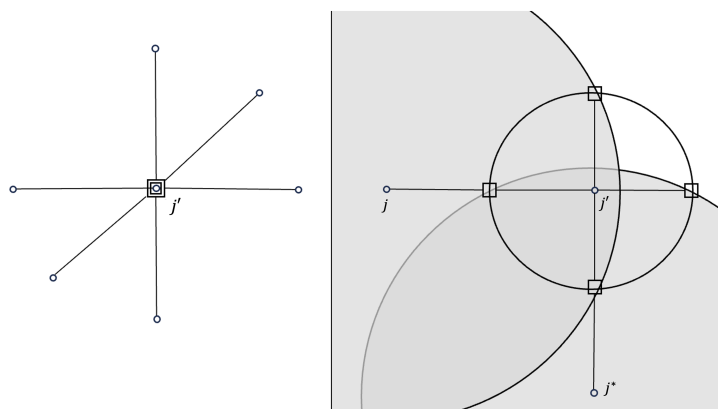
Moreover, **Saving** and **Spending** of center j' is defined as

$$\text{Saving}(j') = \sum_{j \in N_{j'}^+} \{((1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j) - \text{cost}_{j'}(j)\},$$

$$\text{Spending}(j') = \sum_{j \in N_{j'}^-} \{\text{cost}_{j'}(j) - ((1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j)\}.$$

With the goal (1) in mind, $N_{j'}^+$ (resp. $N_{j'}^-$) contains clients j who meet (resp. do not meet) this goal, and $\text{Saving}(j')$ (resp. $\text{Spending}(j')$) indicates how much $\text{cost}_{j'}(j)$ exceeds (resp. is short of) this goal. Note that $N_{j'}^- \subseteq N_{j'} \subseteq N_{j'}^+ \cup N_{j'}^-$ by definition; it is the *best* cluster for center j' , which contains all its neighbors (which is required by the algorithm design) and possibly more to increase savings. Therefore, if $\text{Saving}(j') \geq \text{Spending}(j')$, then j' is considered a *good* center; otherwise, it is a *bad* center.

Now, we will explain some problematic situations that arise when extending the problem to the general case, i.e., without restrictions on C_j , M_j , and D_j 's. The first simple concern is that the choice of the center j' will no longer be solely based on $C_{j'} + M_{j'}$ as in Algorithm 1,



■ **Figure 3** Two examples in a homogeneous network where every neighbor of center j' belongs to $N_{j'}^-$ without creating a dense region.

which breaks the previous arguments. Therefore, to gain more flexibility in selecting a new cluster center, it is beneficial to decompose the entire support graph into several layers, where each layer only concerns clients with roughly the same $C_{j'} + M_{j'}$ values.

► **Definition 5.** A *network* is a subgraph $((\mathcal{F}', \mathcal{C}'), E')$ of the support graph $G = ((\mathcal{F}, \mathcal{C}), E)$. A network is called **homogeneous** if there exists $s \geq 0$, such that for any client $j \in \mathcal{C}'$, $s \leq C_j + M_j \leq (1 + \delta)s$ for $\delta = 3 \times 10^{-23}$.

However, there are still two more scenarios where the above strategy might not hold, as illustrated in Figure 3. This means that the neighbors in $N_{j'}^-$ are all bad clients, but do not create a dense region.

- I. If $C_{j'} \ll s$, the facilities in $\mathcal{C}_{j'}$ are concentrated near j' . In this case, since all the facilities are very close to j' , the 3-hop triangle inequality is almost tight for any $j \in N_{j'}$ regardless of where it is.
- II. Recall that if $\mathcal{C}_j \cup \mathcal{D}_j$ is large enough to exclude the facilities of $\mathcal{C}_{j'}$, then $\text{cost}_{j'}(j)$ might not behave as expected. In particular, a technical problem arises when two facilities from $\mathcal{C}_{j'}$ that are almost antipodal with respect to j' are both contained in $\mathcal{C}_j \cup \mathcal{D}_j$, which is illustrated in Figure 3 (right).

The following definition addresses Scenario I.

► **Definition 6.** For $K_6 = 1.302$, let $\theta = \frac{K_6 + 1 - \gamma}{2K_6 + 2 - \gamma}$. A client $j \in \mathcal{C}$ is **normal** if $C_j \geq \theta(C_j + M_j)$, otherwise **weird**.

The following definition addresses Scenario II. Note that v_j^* is the maximum distance between j and any facility in $\mathcal{C}_j \cup \mathcal{D}_j$. We are interested in the ball around j' of radius $0.998z_{j'}(j)$ (for sake of analysis), and j having a *small remote arm* with respect to j' means that $\mathcal{C}_j \cup \mathcal{D}_j$ cannot contain two antipodal points in our ball of interest (with some slack depending on α). Let $z_{j'}(j) = d(j', \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j))$. The right-hand side represents the square of the length of the other side of a triangle, where the lengths of the two sides are $d(j', j)$ and $0.998z_{j'}(j)$, and the included angle between them is $\frac{\pi}{2} - \alpha$.

► **Definition 7.** For normal center j' , a neighbor j of j' is said to have a **small remote arm** if the following holds for $\alpha = 5 \times 10^{-4}$.

$$(v_j^*)^2 = (C_j^* + F_j^*)^2 < d(j', j)^2 + (0.998z_{j'}(j))^2 - 2d(j', j) \cdot 0.998z_{j'}(j) \cdot \cos\left(\frac{\pi}{2} - \alpha\right)$$

Otherwise, j is said to have a **big remote arm**.

In Section 3.1, we will show that these two scenarios are the only bad cases to worry about. For instance, if we assume that everything is normal and has a small remote arm, each candidate center $j' \in \mathcal{C}$ is either good or contains another candidate center j'' with $Saving(j'') > Saving(j')$ in its dense region.

How can we handle these two bad scenarios? In the following two lemmas, we prove that both the weird center and the big remote arm neighbor imply a high ratio of fractional facility cost F_j^* to fractional connection cost $C_{j'}$. The proof appears in the extended version of this paper [14].

► **Lemma 8.** *For any weird client j , $C_j^* \leq K_6 F_j^*$ holds.*

► **Lemma 9.** *When $1.6 < \gamma < 2$, for a normal center j' , if $j \in N_{j'}^-$ has a big remote arm, then $C_j^* \leq K_6 F_j^*$.*

Therefore, if we consider a (sub)instance that has a low facility-to-connection cost ratio, it is natural to expect to apply this argument, which ultimately leads to a (somewhat) good center. Throughout the paper, we will often express this low facility-to-connection ratio condition will be expressed as $C^* > K F^*$ and use the following values for K : $K_1 = 1.3025 > K_2 = 1.3024 > K_3 = 1.3023 > K_4 = 1.3022 > K_5 = 1.3021 > K_6 = 1.302$. The following theorem is the main result of the section.

► **Theorem 10.** *Consider a homogeneous network and let j' be the normal center with the highest saving among all normal centers. Let $c^* = \sum_{j \in N_{j'}^+ \cup N_{j'}^-} C_j^*$ and $f^* = \sum_{j \in N_{j'}^+ \cup N_{j'}^-} F_j^*$. If $c^* > K_5 f^*$, then this cluster is **good on average** for $\varepsilon_2 = 5 \times 10^{-18}$ and every $\gamma \in (1.6, 2)$. i.e.*

$$\sum_{j \in N_{j'}^+ \cup N_{j'}^-} cost_{j'}(j) \leq \sum_{j \in N_{j'}^+ \cup N_{j'}^-} ((1 - \varepsilon_2)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j).$$

3.1 Geometric Arguments

In this subsection, we prove Theorem 10 using properties of Euclidean geometry. As previously discussed, our goal is to show: When a bad cluster center j' is normal and (many of) its neighbors have a small remote arm, it is possible to find a dense region of clients near j' .

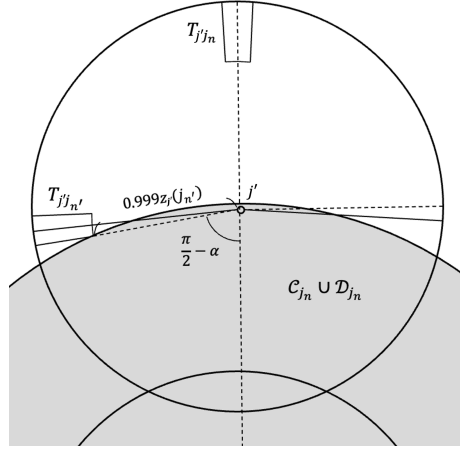
When we consider the rerouting of j to j' 's close facilities $\mathcal{C}_{j'}$, we can define some *worst* facilities. In the below definition, $B_{j',j} \subseteq \mathcal{C}_{j'}$ is the set of facilities that have the (almost) worst distance from j' , and $T_{j',j} \subseteq B_{j',j}$ is the set of facilities with both worst distance and worst angle.

► **Definition 11.** *Let $r := 10^{-8}$, and let $\phi_r \approx 2 \times 10^{-4}$ be the minimum angle that satisfies $1 + x^2 + 2x \cos \phi_r \leq (1 + (1 - r)x)^2$ for $0 \leq x \leq 2$. For $j \in \mathcal{C}$ and its center j' , let*

$$\begin{aligned} B_{j',j} &:= \{i \in \mathcal{C}_{j'} \mid 0.999z_{j'}(j) \leq d(i, j') \leq M_{j'}\} \\ T_{j',j} &:= \{i \in B_{j',j} \mid \angle jj'i > \pi - \phi_r\}. \end{aligned}$$

The following lemma shows that if j' is normal and $j \in N_{j'}^-$ has a bad rerouting through j' , then among the facilities in $B_{j',j} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)$, which are the rerouting candidates with worst distance, more than half of them must have a bad angle as well. It is a formalization of the intuition illustrated in Figure 1.

► **Lemma 12.** *For any $j \in N_{j'}^-$ of normal center j' , $z_{j'}(j) > 0.99C_{j'}$.*



■ **Figure 4** Intersection between a $T_{j'j}$ and a remote arm $C_j \cup D_j$.

Proof. By the triangle inequality and the homogeneous condition,

$$\begin{aligned} z_{j'}(j) &\geq \text{cost}_{j'}(j) - d(j', j) > (1 - \varepsilon_1)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j - (M_{j'} + M_j) \\ &\geq (1 - \varepsilon_1)C_j + (2 - \gamma)M_j + (\gamma - 1)D_j - ((1 + \delta)s - C_{j'}) \\ &\geq -(\varepsilon_1 + \delta)s + C_{j'} \geq \left(1 - \frac{\varepsilon_1 + \delta}{\theta}\right) C_{j'} \geq 0.99C_{j'}. \quad \blacktriangleleft \end{aligned}$$

► **Lemma 13.** For a normal center j' and any $j \in N_{j'}^-$, let $G_1 = T_{j'j} \cap (C_{j'} \setminus (C_j \cup D_j))$, $G_2 = (B_{j'j} \setminus T_{j'j}) \cap (C_{j'} \setminus (C_j \cup D_j))$. Then the following holds:

$$\sum_{i \in G_1} y_i^* > \sum_{i \in G_2} y_i^*.$$

Proof. Assume the nontrivial case: $C_{j'} \setminus (C_j \cup D_j) \neq \emptyset$. Let $G_3 = (C_{j'} \setminus (C_j \cup D_j)) \setminus (G_1 \cup G_2)$. Let rerouting probability p_n and rerouting length l_n for $1 \leq n \leq 3$ as:

$$p_n = \frac{\sum_{i \in G_n} y_i^*}{\sum_{i \in C_{j'} \setminus (C_j \cup D_j)} y_i^*}, \quad l_n = \frac{\sum_{i \in G_n} y_i^* \cdot d(i, j')}{\sum_{i \in G_n} y_i^*}$$

Note that $p_1 + p_2 + p_3 = 1$ and $p_1 l_1 + p_2 l_2 + p_3 l_3 = z_{j'}(j)$. Then the goal of this theorem can be written as $p_1 > p_2$.

By Lemma 12, $z_{j'}(j) \geq 0.99C_{j'}$ holds. We derive a lower bound for $p_1 + p_2$ as:

$$z_{j'}(j) = p_1 l_1 + p_2 l_2 + p_3 l_3 \leq (p_1 + p_2) \cdot M_{j'} + (1 - p_1 - p_2) \cdot 0.999z_{j'}(j)$$

implies that

$$p_1 + p_2 \geq \frac{0.001z_{j'}(j)}{M_{j'} - 0.999z_{j'}(j)} \geq \frac{0.001 \cdot 0.99C_{j'}}{M_{j'}} \geq \frac{0.00099\theta}{1 - \theta}$$

since the right-hand side gives the minimum value when $M_{j'}/C_{j'}$ is the maximum. It is bounded because j' is normal.

70:10 Facility Location on High-Dimensional Euclidean Spaces

Denote a position vector as \vec{v} . From the definition of ϕ_r , $cost_{j'}(j)$ is at most

$$\begin{aligned} cost_{j'}(j) &= \frac{\sum_{i \in \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)} y_i^* \|(v_{j'} - v_j) + (v_i - v_{j'})\|}{\sum_{i \in \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)} y_i^*} \\ &\leq p_1(d(j', j) + l_1) + p_2 \cdot (d(j', j) + (1 - r)l_2) + p_3(d(j', j) + l_3) \\ &= d(j', j) + z_{j'}(j) - p_2 r l_2 \\ &\leq C_{j'} + M_{j'} + (2 - \gamma)M_j + (\gamma - 1)D_j - p_2 r l_2 \\ &\leq (1 - 0.999 \cdot 0.99 p_2 r)C_{j'} + M_{j'} + (2 - \gamma)M_j + (\gamma - 1)D_j. \end{aligned}$$

Therefore, since $j \in N_{j'}^-$, $(1 - \varepsilon_1)C_j + M_j < (1 - 0.999 \cdot 0.99 p_2 r)C_{j'} + M_{j'} \leq (1 - 0.98 p_2 r)C_{j'} + M_{j'}$ holds. It implies

$$\begin{aligned} \left(1 - \frac{\varepsilon_1}{2}\right) (C_j + M_j) &\leq (1 - \varepsilon_1)C_j + M_j < (1 - 0.98 p_2 r)C_{j'} + M_{j'} \\ &\leq ((1 - 0.98 p_2 r)\theta + (1 - \theta))(C_{j'} + M_{j'}). \end{aligned}$$

From the homogeneous condition,

$$\left(1 - \frac{\varepsilon_1}{2}\right) < (1 - 0.98 p_2 r \theta)(1 + \delta)$$

which implies

$$p_2 < \frac{1}{0.98 \theta r} \cdot \frac{\delta + \frac{\varepsilon_1}{2}}{1 + \delta} < \frac{0.00099 \theta}{1 - \theta} < p_1. \quad \blacktriangleleft$$

The following argument demonstrates how the positional distribution of facilities in $\mathcal{C}_{j'}$ restricts that of the neighbors in $N_{j'}^-$. Consider two neighbors $j, k \in N_{j'}^-$, both with small remote arms, separated by an angle greater than $2\phi_r$, say $\frac{1}{100}$. Then, $T_{j'j} \cap T_{j'k} = \emptyset$. However, facilities in $T_{j'j}$ reduce $cost_{j'}(k)$ and vice versa, which implies that either $(\mathcal{C}_k \cup \mathcal{D}_k) \cap T_{j'j} \neq \emptyset$ or vice versa. As the small remote arm condition of j puts a limit on how much $(\mathcal{C}_j \cup \mathcal{D}_j)$ can intersect $\mathcal{C}_{j'}$, it is natural to expect that the number of such pairs is small.

► **Theorem 14.** *For a normal center j' , let S be a subset of $N_{j'}^-$, consisting of clients with a small remote arm. Furthermore, let any two elements $j_1, j_2 \in S$ be separated by an angle greater than $\frac{1}{100}$ with respect to center j' , i.e., $\angle j_1 j' j_2 > \frac{1}{100}$. Then, the cardinality of S is bounded by $M = 5 \times 10^6$, independent of the Euclidean space's dimension.*

Proof. Denote $S = \{j_1, j_2, \dots, j_{|S|}\}$. Without loss of generality,

$$\sum_{i \in T_{j'j_k} \cap (\mathcal{C}_{j'} \setminus (\mathcal{C}_{j_k} \cup \mathcal{D}_{j_k}))} y_i^* \geq \sum_{i \in T_{j'j_{k+1}} \cap (\mathcal{C}_{j'} \setminus (\mathcal{C}_{j_{k+1}} \cup \mathcal{D}_{j_{k+1}}))} y_i^*$$

for all $1 \leq k < |S|$. Suppose $T_{j'j_{n'}} \cap (\mathcal{C}_{j_n} \cup \mathcal{D}_{j_n}) = \emptyset$ for some $n' < n$. Additionally, given $j_n \in N_{j'}^-$ and the small arm condition, it implies $T_{j'j_n} \cap (\mathcal{C}_{j_n} \cup \mathcal{D}_{j_n}) = \emptyset$. Therefore, $T_{j'j_{n'}} \cup T_{j'j_n} \subseteq \mathcal{C}_{j'} \setminus (\mathcal{C}_{j_n} \cup \mathcal{D}_{j_n})$. Moreover, $T_{j'j_{n'}} \cap T_{j'j_n} = \emptyset$ since $2\phi_r < \frac{1}{100}$. However, it contradicts to Lemma 13.

We show that two j 's in S with similar $z_{j'}(j)$ values form a large angle with j' . Suppose $\angle j_{n'} j' j_n = \beta < \frac{\pi}{2} + \alpha - \phi_r$ for some $n' < n$ with $\frac{z_{j'}(j_n)}{z_{j'}(j_{n'})} \in [\frac{1}{1.001}, 1.001]$. Then for any point $x \in T_{j'j_{n'}}$, it holds that $d(j', x) \geq 0.999 z_{j'}(j_{n'}) > 0.998 z_{j'}(j_n)$ and $\angle j_n j' x < \beta + \phi_r < \frac{\pi}{2} + \alpha$. Also, the quadratic function $d(j', x)^2 - 2d(j', j_n)d(j', x) \cdot \sin(\alpha)$ is shown to be non-decreasing

for $d(j', x) > 0$. It comes from Lemma 12, which ensures that $d(j', x) > 0.998z_{j'}(j_n) \geq 0.998 \cdot 0.99C_{j'} \geq 0.98\theta s \geq 2(1 + \delta) \sin \alpha \cdot s \geq d(j', j_n) \cdot \sin \alpha$. Therefore, since j_n has a small remote arm,

$$\begin{aligned} d(j_n, x)^2 &= d(j', j_n)^2 + d(j', x)^2 - 2d(j', j_n)d(j', x) \cdot \cos \angle j_n j' x \\ &> d(j', j_n)^2 + d(j', x)^2 - 2d(j', j_n)d(j', x) \cdot \sin \alpha \\ &\geq d(j', j_n)^2 + (0.998z_{j'}(j_n))^2 - 2d(j', j_n) \cdot 0.998z_{j'}(j_n) \cdot \sin \alpha \\ &> (C_{j_n}^* + F_{j_n}^*)^2 = v_{j_n}^*{}^2 \end{aligned}$$

which implies $T_{j'j_n} \cap (C_{j_n} \cup D_{j_n}) = \emptyset$, contradicting to the above result. Refer to Figure 4.

According to Rankin [17], the maximum number of disjoint spherical caps, each with an angular radius of $\frac{\pi}{4} + \frac{\alpha - \phi_r}{2}$, is at most $1 + \csc(\alpha - \phi_r)$ in any dimension. Since $0.99C_{j'} \leq z_{j'}(j) \leq M_{j'}$ holds, it is feasible to segment this range into successive subranges such as $[0.99C_{j'}, 1.001 \cdot 0.99C_{j'}]$, $[1.001 \cdot 0.99C_{j'}, 1.001^2 \cdot 0.99C_{j'}]$, ..., up to $[M_{j'}/1.001, M_{j'}]$. Thus each subrange can only contain a finite number of clients with a small remote arm. Moreover, the normal center condition ensures a bounded number of such divisions. Consequently, the cardinality of $|S|$ is at most

$$|S| \leq \left(1 + \frac{1}{\sin(\alpha - \phi_r)}\right) \cdot \frac{\log \frac{M_{j'}}{0.99C_{j'}}}{\log 1.001} \leq \left(1 + \frac{1}{\sin(\alpha - \phi_r)}\right) \cdot \frac{\log \frac{1-\theta}{0.99\theta}}{\log 1.001}. \quad \blacktriangleleft$$

Therefore, if we have a large $S \subseteq N_{j'}^-$ with small remote arms, there must exist a large subset of S whose pairwise angle is small, creating a dense region.

► **Lemma 15.** *For a normal center j' , let S be a set of clients from $N_{j'}^-$ with a small remote arm. If $S \neq \emptyset$, then there exists a client $j \in S$ for which $\text{Saving}(j) \geq \frac{s}{125} \frac{|S|}{M}$.*

Proof. Let S' be a maximal subset of S where any two clients are separated by an angle greater than $\frac{1}{100}$ with respect to the center j' . Then, for any client $j \in S$, there exists a client $z_j \in S'$ such that $\angle j j' z_j \leq \frac{1}{100}$. Therefore, there exists a $z \in S'$ such that the number of clients $j \in S$ for which $z_j = z$ is at least $\frac{|S|}{|S'|}$.

For $1 \leq n \leq 5$, let R_n be a region where

$$R_n = \{x \in \mathbb{R}^l \mid \angle z j' x \leq \frac{1}{100}, \frac{2n-2}{5}(1+\delta)s \leq d(j', x) \leq \frac{2n}{5}(1+\delta)s\}.$$

Therefore, there exists an index k such that $|R_k| \geq \frac{|S|}{5|S'|}$. For any two clients $j_1, j_2 \in R_k$, the distance $d(j_1, j_2)$ is bounded by the sum of their radial and angular differences. Hence, $d(j_1, j_2) \leq \frac{2(1+\delta)s}{5} + \frac{2(1+\delta)s}{50} = \frac{11}{25}(1+\delta)s$. From the triangle inequality and the homogeneous condition,

$$\begin{aligned} \text{cost}_{j_1}(j_2) &\leq d(j_1, j_2) + M_{j_1} \leq \frac{36}{25}(1+\delta)s \leq \frac{3-\varepsilon_1}{2(1+\delta)}s \\ &\leq (1-\varepsilon_1)C_{j_2} + 2M_{j_2} \leq (1-\varepsilon_1)C_{j_2} + (3-\gamma)M_{j_2} + (\gamma-1)D_{j_2} \end{aligned}$$

which implies $j_2 \in N_{j_1}^+$. By Theorem 14, $|S'| \leq M$. Therefore, the saving of any client $j \in R_k$ is at least

$$\begin{aligned} \text{Saving}(j) &\geq \frac{|S|}{5|S'|} \left((1-\varepsilon_1)C_j + (3-\gamma)M_j + (\gamma-1)D_j - \frac{36}{25}(1+\delta)s \right) \\ &\geq \frac{|S|}{5M} \frac{3-72\delta-25\varepsilon_1}{50} s \geq \frac{s}{125} \frac{|S|}{M}. \quad \blacktriangleleft \end{aligned}$$

70:12 Facility Location on High-Dimensional Euclidean Spaces

Given these geometric tools, Theorem 10 follows as the number of big-remote-arm-neighbors of the chosen center j' can be bounded.

► **Theorem 10.** *Consider a homogeneous network and let j' be the normal center with the highest saving among all normal centers. Let $c^* = \sum_{j \in N_{j'}^+ \cup N_{j'}^-} C_j^*$ and $f^* = \sum_{j \in N_{j'}^+ \cup N_{j'}^-} F_j^*$. If $c^* > K_5 f^*$, then this cluster is **good on average** for $\varepsilon_2 = 5 \times 10^{-18}$ and every $\gamma \in (1.6, 2)$. i.e.*

$$\sum_{j \in N_{j'}^+ \cup N_{j'}^-} \text{cost}_{j'}(j) \leq \sum_{j \in N_{j'}^+ \cup N_{j'}^-} ((1 - \varepsilon_2)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j).$$

The proof appears in the extended version of this paper [14].

4 Clustering for Homogeneous Instances

In this section, we present an algorithm that operates on a connection-dominant homogeneous instance, ensuring strictly better performance than a naive greedy clustering strategy. Let $c(i)$ for $i \in \mathcal{C}$ be the center of i when some clustering is given in the context.

► **Theorem 16.** *Suppose a homogeneous network $G = ((\mathcal{F}, \mathcal{C}), E)$ satisfies $C^* > K_4 F^*$. Then, the clustering produced by Algorithm 2 is **good on average**. Precisely, for $\varepsilon_3 = 3 \times 10^{-32}$ and every $\gamma \in (1.6, 2)$, the following holds:*

$$\sum_{j \in \mathcal{C}} \text{cost}_{c(j)}(j) \leq \sum_{j \in \mathcal{C}} ((1 - \varepsilon_3)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j).$$

■ **Algorithm 2** HOMOGENEOUS: Homogeneous Clustering.

Require: Network $G = ((\mathcal{F}, \mathcal{C}), E)$

$L = \{\}$

while $\mathcal{C} \neq \emptyset$ **do**

if normal client exists in \mathcal{C} **then**

$c \leftarrow \arg \max_{j \text{ is normal}} \text{Saving}(j)$

else

$c \leftarrow \arg \min_{j \in \mathcal{C}} (C_j + M_j)$

end if

$L \leftarrow L \cup (c, N_c^+ \cup N_c^-)$

$\mathcal{C} \leftarrow \mathcal{C} \setminus (N_c^+ \cup N_c^-)$

end while

return L

For one of a cluster A made by the algorithm, if the center of A is weird, then the ratio is at most K_6 since all clients within them are weird. If A is not “good”, its connection-facility ratio is at most K_5 by Theorem 10. Therefore, the assumed ratio K_4 in this theorem is larger than these values, implying that a constant proportion of clusters are “good” since they satisfy the conditions of Theorem 10. Therefore, by reducing the ε value by that proportion, the desired result can be obtained.

Proof. Divide \mathcal{C} into three groups:

1. Clients clustered by a normal center, where the ratio of that cluster's connection cost to facility cost is greater than K_5 .
2. Clients clustered by a normal center, where this ratio is at most K_5 .
3. Clients clustered by a weird center.

Let S_n be the set of clients in the n -th group ($1 \leq n \leq 3$). Here, S_1 and S_2 correspond to clusters formed by the **if** condition, and S_3 is formed by the **else** condition. Define the following values:

$$C_n^* = \sum_{j \in S_n} C_j^*, \quad F_n^* = \sum_{j \in S_n} F_j^*$$

By Theorem 10, $\sum_{j \in S_1} \text{cost}_{c(j)}(j) \leq \sum_{j \in S_1} ((1-\varepsilon_2)C_j + (3-\gamma)M_j + (\gamma-1)D_j)$. The rerouting cost for S_2 is only bounded by the homogeneous condition. For a client j , $\text{cost}_{c(j)}(j) \leq C_{c(j)} + M_{c(j)} + (2-\gamma)M_j + (\gamma-1)D_j \leq (1+\delta)(C_j + M_j) + (2-\gamma)M_j + (\gamma-1)D_j$. Lastly, clients in S_3 are clustered through a greedy strategy. Thus, $\text{cost}_{c(j)}(j) \leq C_j + (3-\gamma)M_j + (\gamma-1)D_j$. Note that S_3 consists solely of weird clients, meaning $C_3^* \leq K_6 F_3^* \leq K_5 F_3^*$.

From above, the total rerouting cost is bounded by:

$$\begin{aligned} \sum_{j \in \mathcal{C}} \text{cost}_{c(j)}(j) &= \sum_{j \in S_1} \text{cost}_{c(j)}(j) + \sum_{j \in S_2} \text{cost}_{c(j)}(j) + \sum_{j \in S_3} \text{cost}_{c(j)}(j) \\ &\leq \sum_{j \in S_1} ((1-\varepsilon_2)C_j + M_j + (2-\gamma)M_j + (\gamma-1)D_j) \\ &\quad + \sum_{j \in S_2} ((1+\delta)(C_j + M_j) + (2-\gamma)M_j + (\gamma-1)D_j) \\ &\quad + \sum_{j \in S_3} (C_j + (3-\gamma)M_j + (\gamma-1)D_j). \end{aligned}$$

Therefore, it is sufficient to show that

$$(\delta + \varepsilon_3) \sum_{j \in S_2} (C_j + M_j) + \varepsilon_3 \sum_{j \in S_3} C_j \leq (\varepsilon_2 - \varepsilon_3) \sum_{j \in S_1} C_j.$$

Note that $C_2^* + C_3^* \leq K_5(F_2^* + F_3^*) < \frac{K_5}{K_4} C^*$ holds, which means $C_1^* > \left(1 - \frac{K_5}{K_4}\right) C^*$. Also $C_1^* > K_4 F_1^*$, since $C_2^* + C_3^* \leq K_5(F_2^* + F_3^*)$. Then the following holds:

$$\begin{aligned} &(\delta + \varepsilon_3) \sum_{j \in S_2} (C_j + M_j) + \varepsilon_3 \sum_{j \in S_3} C_j \\ &\leq (\delta + \varepsilon_3) \sum_{j \in S_2 \cup S_3} (C_j + M_j) \leq (\delta + \varepsilon_3) \sum_{j \in \mathcal{C}} (C_j + D_j) \\ &\leq (\delta + \varepsilon_3)(2C^* + (2-\gamma)F^*) \\ &\leq \frac{(\delta + \varepsilon_3)(2-\gamma + 2K_4)}{K_4} C^* \\ &\leq (\varepsilon_2 - \varepsilon_3) \frac{(K_5 - \gamma + 1)(K_4 - K_5)}{K_4 K_5} C^* \leq (\varepsilon_2 - \varepsilon_3) \frac{K_5 - \gamma + 1}{K_5} C_1^* \\ &\leq (\varepsilon_2 - \varepsilon_3)(C_1^* - (\gamma-1)F_1^*) \leq (\varepsilon_2 - \varepsilon_3) \sum_{j \in S_1} C_j. \quad \blacktriangleleft \end{aligned}$$

5 Clustering for Connection-dominant Instances

In this section, we introduce an algorithm that operates on connection-dominant instances without a homogeneous condition. This algorithm uses Algorithm 1 and Algorithm 2 as subroutines. The theorem stated below shows our main result.

► **Theorem 17.** *For any connection-dominant instance, i.e. $C^* > K_1 F^*$, there exists an algorithm that finds a clustering configuration whose rerouting cost is at most*

$$\sum_{j \in \mathcal{C}} \text{cost}_{c(j)} \leq \sum_{j \in \mathcal{C}} ((1 - \varepsilon_5)C_j + (3 - \gamma)M_j + (\gamma - 1)D_j)$$

for $\varepsilon_5 = 2 \times 10^{-41}$ and every $\gamma \in (1.6, 2)$.

► **Definition 18.** *Let B_0 be the set of clients for which $C_j + M_j = 0$. Let $s = \min_{j \in \mathcal{C} \setminus B_0} (C_j + M_j)$. For $\delta' = 7 \times 10^{-32}$, let a **Block** B_n be the set of clients such that*

$$B_n = \{j \in \mathcal{C} \mid (1 + \delta')^{n-1}s \leq C_j + M_j < (1 + \delta')^n s\}$$

Note that if $(1 + \delta')^m \leq (1 + \delta)$, then a network composed of at most m consecutive blocks is still homogeneous. However, applying Algorithm 2 directly to each block would be impossible when the block is a facility-dominant or neighbors of the center from some block may belong to a different block. Moreover, the following fact implies that the neighbor relationships between consecutive blocks are the main point.

► **Observation 19.** *If two neighbors $j, j' \in \mathcal{C}$ belonging to neither the same block nor consecutive blocks, satisfying $(1 + \delta')(C_j + M_j) \leq C_{j'} + M_{j'}$, then $j' \in N_j^+$ holds when *Saving*(\cdot) criteria ε satisfies that $\varepsilon \leq \frac{2\delta'}{1+\delta'}$.*

Proof.

$$\begin{aligned} \text{cost}_j(j') &\leq C_j + M_j + (2 - \gamma)M_{j'} + (\gamma - 1)D_{j'} \leq \frac{C_{j'} + M_{j'}}{1 + \delta'} + (2 - \gamma)M_{j'} + (\gamma - 1)D_{j'} \\ &\leq \left(1 - \frac{2\delta'}{1 + \delta'}\right) \cdot C_{j'} + (3 - \gamma)M_{j'} + (\gamma - 1)D_{j'}. \quad \blacktriangleleft \end{aligned}$$

A key idea is that with a sufficiently small δ' , the support graph can be segmented almost arbitrarily while still preserving the homogeneous condition, enabling the identification of weak *connections* between consecutive blocks. Notably, even if two clients $j \in B_n$ and $j' \in B_{n+1}$ come from consecutive blocks, the $\text{cost}_j(j')$ is not worse than the greedy strategy.

► **Definition 20.** *An **interval** I is the set of consecutive blocks, up to a maximum $2L$ for $L = 2 \times 10^8$. Precisely, $I = \{B_i, \dots, B_{i+k}\}$ for $k < L$, where $\sum_{j=i}^{i+k} C^*(B_j) - C^*(B_i) \geq K_3 \sum_{j=i}^{i+k} F^*(B_j)$, except when $k = 0$. The **size** of an interval I is the number of blocks it contains, denoted as $|I|$. The **reward** of interval I is defined as $R(I) = \sum_{j=i}^{i+k} C^*(B_j) - C^*(B_i)$.*

Note that a block is a unique type of interval with a size of 1. Also $(1 + \delta')^{2L} \leq 1 + \delta$ holds. An interval is the basic unit of clustering to which Algorithm 1 and Algorithm 2 are applied. Therefore, the reward of an interval represents the extent to which a guaranteed *Saving* can be obtained when the current interval can be clustered using Algorithm 2, regardless of how the preceding intervals have been clustered. From this perspective, the first block, which does not contribute to the reward, serves as a kind of “buffer”.

Given the entire support graph $G = (\mathcal{F} \cup \mathcal{C}, E)$ and a subset $\mathcal{C}' \subseteq \mathcal{C}$, let $G_{\mathcal{C}'}$ be the subgraph (network) induced by $\mathcal{F} \cup \mathcal{C}'$. Consequently, when we call Algorithms 1 or 2 with $G_{\mathcal{C}'}$, the calculations for *Saving* and *Spending* are performed solely with respect to the implicitly defined client set \mathcal{C}' . We denote Algorithm 1 as GREEDY and Algorithm 2 as HOMOGENEOUS. For simplicity, when I denotes an interval, we interpret the expression $\sum_{j \in I}$ as $\sum_{B \in I} \sum_{j \in B}$, aggregating over all clients within the interval. The following theorem illustrates how reward is related to *Saving*. The proof appears in the extended version of this paper [14].

► **Lemma 21.** *Let J be a set of non-overlapping intervals. Then clustering produced by Algorithm 3 is **good on average**. Precisely, for $\varepsilon_4 = 2 \times 10^{-36}$,*

$$\sum_{j \in \mathcal{C}} \text{cost}_{c(j)}(j) \leq \sum_{j \in \mathcal{C}} \left(\left(1 - \frac{\sum_{I \in J} R(I)}{\sum_{j' \in \mathcal{C}} C_{j'}} \varepsilon_4 \right) C_j + (3 - \gamma)M_j + (\gamma - 1)D_j \right).$$

■ **Algorithm 3** CONN: Connection-dominant Clustering.

Require: Set of disjoint intervals $J = \{I_1, I_2, \dots, I_m\}$, arranged in the increasing order of $C_j + M_j$ values of their clients.

$L = \{\}$

for $k = 1$ to m **do**

if $|I_k| \geq 2$ and $\sum_{j \in I_k} C_j^* > K_4 \sum_{j \in I_k} F_j^*$ **then**

$L \leftarrow L \cup \text{HOMOGENEOUS}(G_{I_k})$

else

$L \leftarrow L \cup \text{GREEDY}(G_{I_k})$

end if

$S := \bigcup_{j \in I_k} (N_j^+ \cup N_j^-)$

for $l = k$ to m **do**

$I_l \leftarrow I_l \setminus S$

end for

end for

return L

In conclusion, it suffices to find a set of non-overlapping intervals J which have a high $R(J)$ value. Here, we will briefly touch on the idea. We will iterate through the blocks in reverse order (B_N, \dots, B_0) , cutting out suitable ranges that satisfy the interval conditions. Fix a point r to be the right end of the interval, and expand the left end of the current range one block to the left until the range is suitable for processing. Suppose, at some point, the C/F value of the current range is less than K_2 , which is strictly less ratio of the input instance, K_1 . Then, this range can be considered a minor part and can be excluded as there is no need for it to become an interval. Therefore, we only consider cases where the current range has a C/F value of at least K_2 .

However, in the situation where the current range's reward is small, meaning that the first block must avoid most of the C^* . In this case, if we keep expanding the range to the left, we will eventually reach the initial block B_0 , satisfying the interval condition. There remains a subtle issue of the range's length exceeding $2L$ during the expansion process, but this implies that the C^* value on the left side of the range is always exponentially increasing compared to the right side. This can be resolved by appropriately reducing the right side of the range.

► **Lemma 22.** For any support graph G that is connection-dominant, $C^* > K_1 F^*$, then the set of non-overlapping intervals J obtained by Algorithm 4 satisfies $\sum_{I \in J} R(I) \geq \frac{1}{10^5} C^*$.

■ **Algorithm 4** CUTINTERVAL: Find a set of non-overlapping intervals J with large rewards.

Require: Network $G = ((\mathcal{F}, \mathcal{C}), E)$.

$J := \emptyset$

$r := N$

while $r > 0$ **do**

$l := r$

$c := 0, f := 0$

while $l \geq 0$ **do**

if $c \geq K_3(f + F^*(B_l))$ and $C^*(B_l) \leq \frac{K_2 - K_3}{K_2} c$ **then**

$J \leftarrow J \cup \{[l, r]\}$

$r \leftarrow l - 1$

break

end if

$c \leftarrow c + C^*(B_l), f \leftarrow f + F^*(B_l)$

if $c < K_2 f$ **then**

$r \leftarrow l - 1$

break

end if

if $r - l + 1 = 2L$ **then**

$c \leftarrow c - \sum_{i=l+L}^r C^*(B_i), f \leftarrow f - \sum_{i=l+L}^r F^*(B_i)$

$r \leftarrow r - L$

end if

$l \leftarrow l - 1$

end while

end while

Include in J every block (as an interval of size 1) not already part of J

return J

The proof appears in the extended version of this paper [14].

By directly applying Lemma 21 and Lemma 22, Theorem 17 can be proved.

6 Improved Bifactor Approximation

In this section, we present an improved bifactor approximation algorithm for UFL, proving Theorem 1. To deal with facility-dominant instances, we employ the JMS algorithm [12], which is known to be $(1.11, 1.7764)$ -approximation algorithm.

Now, we will show that Algorithm 5 satisfies Theorem 1. Let $\gamma_0 \leq 1.6774$ be a solution of the below equation.

$$\frac{1}{e} + \frac{1}{e^\gamma} - (\gamma - 1)\left(1 - \frac{1}{e} + (1 - \varepsilon_5)\frac{1}{e^\gamma}\right) = 0.$$

To prove Theorem 1, we consider two cases. When $C^* \leq K_1 F^*$, the inequality $1.11F^* + 1.7764C^* \leq 1.6774F^* + (1 + 2e^{-1.6774} - \varepsilon_6)C^*$ holds. For the case $C^* > K_1 F^*$, we can employ the same proof structure as the statement proved in [3]². The proof appears in the extended version of this paper [14].

² Theorem 4.3 of [3]

■ **Algorithm 5** Overall bi-factor clustering process.

Derive (\bar{x}, \bar{y}) from the optimal LP primal solution (x^*, y^*) .
if $C^* \leq K_1 F^*$ **then**
 Execute the (1.11, 1.7764)-approximation algorithm.
else
 Proceed with $\text{CONN}(\text{CUTINTERVAL}(G))$ (Algorithm 3 and 4).
 for all cluster centers j **do**
 Open exactly one nearby facility with a probability of \bar{x}_{ij} .
 end for
 for all facilities i not close to any cluster center **do**
 Open each facility with a probability of \bar{y}_i independently.
 end for
 Connect each client to the nearest open facility.
end if

7 Improved Unifactor Approximation

In this section, we propose the algorithm that guarantees better unifactor approximation suggested by Li [15], proving Theorem 2.

Framework of [15]. Li showed that a hard instance for a certain γ might not be a hard instance for another value of γ . This suggests that selecting a γ value at random could improve the expected performance of the algorithm.

They introduced a *characteristic function* $h : (0, 1] \rightarrow \mathbb{R}$ to represent the distribution of distances between a client and its neighboring facilities. For a client $j \in \mathcal{C}$, assume i_1, i_2, \dots, i_k are the facilities within $\mathcal{C}_j \cup \mathcal{D}_j$, ordered by increasing distance from j . Then, for $0 < p \leq 1$, $h_j(p)$ is defined as $d(j, i_t)$, where t is the smallest index satisfying $\sum_{l=1}^t y_{i_l}^* \geq p$. Also, they improved the bound of the rerouting cost as $d(j, \mathcal{C}_{j'} \setminus (\mathcal{C}_j \cup \mathcal{D}_j)) \leq (2 - \gamma)M_j + (\gamma - 1)D_j + C_{j'} + M_{j'}$. For fixed γ , the expected connection cost for j for the aforementioned algorithm of [3] is at most

$$\mathbb{E}[C_j] \leq \int_0^1 h_j(p) e^{-\gamma p} \gamma dp + e^{-\gamma} \left(\gamma \int_0^1 h_j(p) dp + (3 - \gamma) h_j\left(\frac{1}{\gamma}\right) \right).$$

Therefore, it can be modeled as a 0-sum game to analyze the approximation ratio. The characteristic function for the whole instance is given by $h(p) = \sum_{j \in \mathcal{C}} h_j(p)$. Assuming h is normalized so that $\int_0^1 h(p) dp = 1$, the algorithm proceeds as follows: with probability κ , it employs the JMS algorithm [12]. Mahdian [16] proved that the JMS algorithm achieves a (1.11, 1.7764)-approximation. Otherwise, γ is sampled randomly from the distribution $\mu : (1, \infty] \rightarrow \mathbb{R}^*$, ensuring that $\kappa + \int_1^\infty \mu(\gamma) d\gamma = 1$. Thus, the value of the 0-sum game, i.e., the approximation ratio of the algorithm under a fixed strategy (κ, μ) , is calculated as follows:

$$\nu(\kappa, \mu, h) = \max \left\{ \int_1^\infty \gamma \mu(\gamma) d\gamma + 1.11\kappa, \int_1^\infty \alpha(\gamma, h) \mu(\gamma) d\gamma + 1.7764\kappa \right\}$$

where

$$\alpha(\gamma, h) = \int_0^1 h(p) e^{-\gamma p} \gamma dp + e^{-\gamma} \left(\gamma + (3 - \gamma) h\left(\frac{1}{\gamma}\right) \right).$$

70:18 Facility Location on High-Dimensional Euclidean Spaces

Moreover, for a given probability density function μ for γ , it can be shown that the characteristic function for the hardest instance is a *threshold function*, which defined as

$$h_q(p) = \begin{cases} \frac{1}{1-q}, & \text{for } p > q \\ 0, & \text{for } p \leq q \end{cases}$$

for some $0 \leq q < 1$. This means that the final approximation ratio for some μ is given by $\max_{0 \leq q < 1} \nu(\kappa, \mu, h_q)$. In [15], the suggested distribution for α_{Li} is $\mu(p) = \theta D(p - \gamma_1) + \frac{1-\kappa-\theta}{\gamma_2-\gamma_1} [\gamma_1 < p < \gamma_2]$, where D is Dirac delta function, $\gamma_1 = 1.479311$, $\gamma_2 = 2.016569$, $\theta = 0.503357$, $\kappa = 0.195583$.

Our Improvement. By using Algorithm 5, the expected connection cost for some γ and characteristic function h is given as follows:

$$\alpha'(\gamma, h) = \begin{cases} \int_0^1 h(p) e^{-\gamma p} dp + e^{-\gamma} \left(\gamma(1 - \varepsilon_7) + (3 - \gamma)h\left(\frac{1}{\gamma}\right) \right), & \text{if } 1.6 < \gamma < 2 \\ \int_0^1 h(p) e^{-\gamma p} dp + e^{-\gamma} \left(\gamma + (3 - \gamma)h\left(\frac{1}{\gamma}\right) \right). & \text{otherwise} \end{cases}$$

Therefore, we define a new 0-sum game value as follows, assuming h is scaled up such that $\int_0^1 h(p) dp = 1$.

$$\nu'(\mu, \theta, h) = \max \left\{ \int_1^\infty \gamma \mu(\gamma) d\gamma + 1.11\kappa, \int_1^\infty \alpha'(\gamma, h) \mu(\gamma) d\gamma + 1.7764\kappa \right\}.$$

Note that α' is still linear for h . Therefore, even if the game definition changes, the adversary's choice of the characteristic function h remains a threshold function h_q for some $0 \leq q < 1$. Given that there is a positive probability of sampling γ between 1.6 and 2, it is possible to achieve a lower cost.

► **Lemma 23.** *Let $\mu_2(\gamma) = (1 - \varepsilon_7)\mu_1(\gamma) + \varepsilon_7(1 - \kappa_2)D(\gamma - 1)$, where D is a Dirac-delta function. Then the following holds:*

$$\begin{aligned} & \max_{0 \leq q < 1} \max \left\{ \int_1^\infty \gamma \mu_2(\gamma) d\gamma + 1.11\kappa_2, \int_1^\infty \alpha'(\gamma, h_q) \mu_2(\gamma) d\gamma + 1.7764\kappa_2 \right\} \\ & < \max_{0 \leq q < 1} \max \left\{ \int_1^\infty \gamma \mu_1(\gamma) d\gamma + 1.11\kappa_2, \int_1^\infty \alpha(\gamma, h_q) \mu_1(\gamma) d\gamma + 1.7764\kappa_2 \right\} - \frac{\varepsilon_7}{1000}. \end{aligned}$$

The proof appears in the extended version of this paper [14].

Therefore, we present an improved unifactor approximation algorithm, Algorithm 6.

When $C^* \leq K_1 F^*$, an inequality $1.11F^* + 1.7764C^* \leq 1.487F^* + 1.487C^*$ is satisfied. For the case $C^* > K_1 F^*$, by Lemma 23, Algorithm 6 shows an improved unifactor approximation with $\varepsilon_8 = \frac{\varepsilon_7}{1000} = 2 \times 10^{-45}$.

8 APX-Hardness

In this section, we prove that UFL is APX-Hard in Euclidean spaces. We use the following result from Austrin, Khot, and Safra [2]. For $\rho \in [-1, +1]$ and $\mu \in [0, 1]$, let $\Gamma_\rho(\mu) := \Pr[X \leq \Phi^{-1}(\mu) \wedge Y \leq \Phi^{-1}(\mu)]$ where X, Y are standard Gaussian random variables with covariance ρ and Φ is the cumulative density function of the standard normal distribution.

■ **Algorithm 6** Overall uni-factor clustering process.

Derive (\bar{x}, \bar{y}) from the optimal LP primal solution (x^*, y^*) .
 Let X be a random variable that takes the value 1 with probability $1 - \kappa_2$, and 0 otherwise.

if $C^* \leq K_1 F^*$ **then**
 Execute the (1.11, 1.7764)-approximation algorithm.

else if $X = 1$ **then**
 Execute the (1.11, 1.7764)-approximation algorithm.

else
 Takes γ from the distribution $\mu_2(\gamma) = (1 - \varepsilon_\gamma)\mu_1(\gamma) + \varepsilon_\gamma(1 - \kappa_2)D(\gamma - 1)$.
 if $1.6 \leq \gamma \leq 2$ **then**
 Proceed with $\text{CONN}(\text{CUTINTERVAL}(G))$ (Algorithm 3 and 4).
 else
 Proceed with $\text{GREEDY}(G)$.
 end if
 for all cluster centers j **do**
 Open exactly one nearby facility with a probability of \bar{x}_{ij} .
 end for
 for all facilities i not close to any cluster center **do**
 Open each facility with a probability of \bar{y}_i independently.
 end for
 Connect each client to the nearest open facility.
end if

► **Theorem 24.** *Assuming the Unique Games Conjecture, for any $q \in (0, 1/2)$ and $\varepsilon > 0$, it is NP-hard to, given a graph $G = (V, E)$, distinguish between the following two cases.*

- (Completeness) G contains an independent set of size $q|V|$.
- (Soundness) For any $T \subseteq V$, the number of edges with both endpoints in T is at least $|E| \cdot (\Gamma_{-q/(1-q)}(\mu) - \varepsilon)$ where $\mu = |T|/|V|$.

Fix an arbitrary $q \in (0, 1/2)$. Without loss of generality, assume $V = [n]$. Also let $m := |E|$. Our UFL instance has V as the set of facilities and E as the set of clients. The ambient Euclidean space is \mathbb{R}^n , and let e_i be the i th standard unit vector (i.e., $(e_i)_i = 1$ and $(e_i)_j = 0$ for every $j \neq i$). Then each facility $i \in V$ is located at e_i and each client $(i, j) \in E$ is located at $e_i + e_j$. Finally, let λ be the common facility cost for every $i \in \mathcal{F}$ to be determined. This finishes the description of the UFL instance.

In the completeness case, there is an independent set U of size qn . We open $V \setminus U$. Since $V \setminus U$ is a vertex cover, every client in E has a facility at distance 1, so the total cost is

$$\lambda(1 - q)n + m.$$

In the soundness case, consider any solution that opens $S \subseteq V$ and let $T = V \setminus S$ and $\mu = |T|/n$. By the soundness guarantee, at least $m(\Gamma_{-q/(1-q)}(\mu) - \varepsilon)$ clients do not have a facility at distance 1. Since every client-facility distance is either 1 or $\sqrt{3}$, the total cost is at least

$$\lambda(1 - \mu)n + m(1 + (\sqrt{3} - 1)(\Gamma_{-q/(1-q)}(\mu) - \varepsilon)). \quad (2)$$

For fixed q , the function $\Gamma_{-q/(1-q)}(\mu)$ is a strictly convex function of μ , so if we let λ such that

$$-\lambda n + m(\sqrt{3} - 1) \frac{d\Gamma_{-q/(1-q)}(\mu)}{d\mu} \Big|_{\mu=q} = 0,$$

then (2) is minimized when $\mu = q$, which becomes

$$\lambda(1 - q)n + m(1 + (\sqrt{3} - 1)(\Gamma_{-q/(1-q)}(q) - \varepsilon)).$$

Furthermore, we can notice that

$$\lambda = \frac{m}{n} \cdot (\sqrt{3} - 1) \frac{d\Gamma_{-q/(1-q)}(\mu)}{d\mu} \Big|_{\mu=q} = \Theta\left(\frac{m}{n}\right).$$

Then one can see that the optimal value in the soundness case is at least $m(\sqrt{3} - 1)(\Gamma_{-q/(1-q)}(q) - \varepsilon)$ larger than the optimal value in the completeness case. For a fixed q , by choosing ε sufficiently small, one can ensure that this excess is at least a δ fraction of the completeness case optimal value for some constant $\delta > 0$, which proves a $(1 + \delta)$ -hardness of approximation.

9 Conclusion

The most natural open problem is to get an improved approximation for bifactor or unifactor approximation for UFL. Though we show a strict separation between general and Euclidean metrics for bifactor approximation in a certain regime, it is not achieved for all regimes of bifactor or unifactor approximation.

Whereas our algorithm is based on the *primal rounding* approach of [3] and [15], it might be a fruitful research direction to design a variant of the Jain-Mahdian-Saberi algorithm [12] (greedy algorithm analyzed by the dual fitting method) or Jain-Vazirani [13] (primal-dual algorithm) for a further improvement. In particular, as the best unifactor approximation for UFL in both general and Euclidean metrics employ the $(1.11, 1.7764)$ -approximation of the JMS algorithm as a black box, improving the JMS algorithm will directly yield a better result for the best unifactor approximation for UFL. The JV algorithm was already improved in Euclidean spaces [1, 10, 4], but they are not enough for UFL.

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. *SIAM Journal on Computing*, 49(4):FOCS17–97, 2019. doi:10.1137/18M1171321.
- 2 Per Austrin, Subhash Khot, and Muli Safra. Inapproximability of vertex cover and independent set in bounded degree graphs. *Theory of Computing*, 7(1):27–43, 2011. doi:10.4086/TOC.2011.V007A003.
- 3 Jaroslaw Byrka and Karen Aardal. An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem. *SIAM Journal on Computing*, 39(6):2212–2231, 2010. doi:10.1137/070708901.
- 4 Vincent Cohen-Addad, Hossein Esfandiari, Vahab Mirrokni, and Shyam Narayanan. Improved approximations for euclidean k-means and k-median, via nested quasi-independent sets. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1621–1628, 2022. doi:10.1145/3519935.3520011.

- 5 Vincent Cohen-Addad, Andreas Emil Feldmann, and David Saulpic. Near-linear time approximation schemes for clustering in doubling metrics. *Journal of the ACM (JACM)*, 68(6):1–34, 2021. doi:10.1145/3477541.
- 6 Vincent Cohen-Addad and CS Karthik. Inapproximability of clustering in ℓ_p metrics. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 519–539. IEEE, 2019.
- 7 Vincent Cohen-Addad, Karthik C S, and Euiwoong Lee. Johnson coverage hypothesis: Inapproximability of k-means and k-median in ℓ_p -metrics. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1493–1530. SIAM, 2022. doi:10.1137/1.9781611977073.63.
- 8 Vincent Cohen-Addad Viallat, Fabrizio Grandoni, Euiwoong Lee, and Chris Schwiegelshohn. Breaching the 2 $\ln p$ approximation barrier for facility location with applications to k-median. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 940–986. SIAM, 2023.
- 9 Kishen N Gowda, Thomas Pensyl, Aravind Srinivasan, and Khoa Trinh. Improved bi-point rounding algorithms and a golden barrier for k-median. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 987–1011. SIAM, 2023. doi:10.1137/1.9781611977554.CH38.
- 10 Fabrizio Grandoni, Rafail Ostrovsky, Yuval Rabani, Leonard J Schulman, and Rakesh Venkat. A refined approximation for euclidean k-means. *Information Processing Letters*, 176:106251, 2022. doi:10.1016/J.IPL.2022.106251.
- 11 Sudipto Guha and Samir Khuller. Greedy strikes back: Improved facility location algorithms. *Journal of algorithms*, 31(1):228–248, 1999. doi:10.1006/JAGM.1998.0993.
- 12 Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pages 731–740, 2002. doi:10.1145/509907.510012.
- 13 Kamal Jain and Vijay V Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and lagrangian relaxation. *Journal of the ACM (JACM)*, 48(2):274–296, 2001. doi:10.1145/375827.375845.
- 14 Euiwoong Lee and Kijun Shin. Facility location on high-dimensional euclidean spaces, 2024. arXiv:2501.18105.
- 15 Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Information and Computation*, 222:45–58, 2013. doi:10.1016/J.IC.2012.01.007.
- 16 Mohammad Mahdian, Yinyu Ye, and Jiawei Zhang. A 1.52-approximation algorithm for the uncapacitated facility location problem. In *Proc. of APPROX*, pages 229–242, 2002.
- 17 Robert Alexander Rankin. The closest packing of spherical caps in n dimensions. *Glasgow Mathematical Journal*, 2(3):139–144, 1955.