Approximating Lower- and Upper-Bounded (Connected) Facility Location via LP Rounding

Zachary Friggstad*, Mohsen Rezapour, and Mohammad R. Salavatipour†

Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada, T6G 2E8
{zacharyf,rezapour,mrs}@ualberta.ca

Abstract
We consider a lower- and upper-bounded generalization of the classical facility location problem, where each facility has a capacity (upper bound) that limits the number of clients it can serve and a lower bound on the number of clients it must serve if it is opened. We develop an LP rounding framework that exploits a Voronoi diagram-based clustering approach to derive the first bicriteria constant approximation algorithm for this problem with non-uniform lower bounds and uniform upper bounds. This naturally leads to the first LP-based approximation algorithm for the lower bounded facility location problem (with non-uniform lower bounds).

We also demonstrate the versatility of our framework by extending this and presenting the first constant approximation algorithm for some connected variant of the problems in which the facilities are required to be connected as well.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Facility Location; Approximation Algorithm; LP Rounding

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

We study the lower- and upper-bounded facility location (LUFL) problem, a natural generalization of the well-known capacitated facility location (CFL) and lower bounded facility location (LBFL) problems. We are given a complete graph $G = (V,E)$, with metric edge lengths $c_e \in \mathbb{Z}_{\geq 0}, e \in E$ containing a set of potential facilities $F \subseteq V$ and a set of demand points (clients) $D \subseteq V$. Each facility $i \in F$ has an opening cost $\mu_i \in \mathbb{Z}_{\geq 0}$ and a capacity (upper bound) $U_i \in \mathbb{Z}_{>0}$, which limits the amount of demand it can serve. Moreover, each facility $i$ has a lower bound $L_i \in \mathbb{Z}_{\geq 0}$ on the amount of demand it must serve if it is opened.

A feasible solution to LUFL consists of a set of facilities $I \subseteq F$ to open, and a valid assignment $\sigma : D \rightarrow I$ of clients to the open facilities: an assignment is valid if it satisfies the lower and upper bounds

$$L_i \leq |\sigma^{-1}(i)| \leq U_i \quad \forall i \in I.$$ (1)

The goal is to minimize the total cost, i.e., $\sum_{i \in I} \mu_i + \sum_{j \in D} c_{\sigma(j)}$.

In many real-world applications, particularly in telecommunications, there is an additional requirement to connect the open facilities via high bandwidth core cables. This leads to a variant of LUFL in which open facilities are connected via a tree-like core network that

* Supported by NSERC and funding from the Canada Research Chairs program.
† Supported by NSERC.
consists of infinite capacity cables. We model this variant as a connected lower- and upper-bounded facility location (C-LUFL) problem. Let us introduce a parameter $M \geq 1$ which reflects the cost per unit length of core cables. A feasible solution to C-LUFL is given by a set of facilities $I \subseteq F$, an assignment $\sigma : D \rightarrow I$ of clients to the open facilities that is valid, and a Steiner tree of $T \subseteq E$ connecting all facilities $I$ via core cables. The objective of C-LUFL is to minimize the total cost, i.e., $\sum_{i \in I} \mu_i + \sum_{j \in D} c_{\sigma(j)} + M \sum_{e \in T} c_e$.

Both the CFL and LBFL problems have been well-studied in the literature. However, there is not much work in studying these problems in a complementary way. To address this gap of knowledge, in this paper, we develop a framework that combines LP rounding techniques for facility location problems with a Voronoi diagram-based clustering approach in order to obtain the first (bicriteria) approximation algorithms for several variants of the problems.

**Definition 1.** An $(\rho, \alpha, \beta)$-approximation algorithm for LUFL (C-LUFL, resp.) is an algorithm that computes in polynomial time a solution $(I, \sigma)$ satisfying $\left\lfloor \frac{L_i}{\alpha} \right\rfloor \leq |\sigma^{-1}(i)| \leq \left\lceil \beta U_i \right\rceil$, $\forall i \in I$, with cost at most $\rho \cdot OPT$, where OPT denotes the minimum cost of a solution to LUFL (C-LUFL, resp.) satisfying (1).

We often loosely refer to a $(\rho, \alpha, \beta)$-approximation for LUFL or C-LUFL when $\rho, \alpha, \beta$ are constants as a relaxed constant-factor approximation.

**Related Work.** The CFL problem is the special case of LUFL when $L_i = 0$ for all $i \in F$. There are several approximation algorithms for CFL based on local search techniques. For the case of uniform capacities, Korupolu et al. [13] gave the first constant factor approximation algorithm, with ratio 8. This was later improved to 5.83 [6] and 3 [2]. The first constant factor approximation for the case of non-uniform capacities was proposed by [17] who gave an 9-approximation, which was eventually improved to 5 [5]. An LP-based approach to CFL was employed by Shmoys et al. [18] who gave the first bicriteria approximation for uniform capacities; this was extended to non-uniform capacities [1]. Levi et al. [14] obtained an LP-based 5-approximation algorithm when facilities opening costs are uniform. For a long time it was an open question to prove a constant factor approximation for CFL based on LP-rounding. This was recently answered by An et al. [4] who gave an LP-based 288-approximation algorithm for CFL which works for the general case.

The LBFL problem is another special cases of LUFL when $U_i = \infty$ for all $i \in F$. This problem was introduced independently by Guha et al. [9] and Karger et al. [12] who gave a bicriteria approximation. The first true approximation algorithm for LBFL was given by Svitkina [19] with ratio 448. The factor was then improved to 82.6 by [3] by applying a modified variant of the algorithm of [19], combined with a more careful analysis. We note that the approaches of both papers work only if all lower bounds are uniform. Finding a true approximation for LBFL when the lower bounds are non-uniform remains an open problem. To the best of our knowledge, there have been no LP-based approximation (even bicriteria) algorithms for LBFL in the literature.

The Connected Facility Location (ConFL) problem is an obvious special case of C-LUFL (when $U_i = \infty$ & $L_i = 0$ for all $i \in F$.). The ConFL problem was first introduced by Gupta et al. [10], in the context of reserving bandwidth for virtual private networks, where they gave the first constant-factor approximation algorithm for ConFL. Using the

---

1 In an earlier version of [1] there was an attempt to study LUFL but there seemed to be an error in the proof. After checking with the authors the claim about LUFL is retracted in the current version of [1].
primal-dual technique, the factor was then improved to 8.55 by [20], and to 6.55 by [11]. Applying sampling techniques, the guarantee was later reduced to 4 by [7], and to 3.19 by [8].

**Our Results and Techniques.** We explore LP-based approaches to obtain bicriteria approximations for many combinations of lower/upper/connected facility location. Our first main result is the first constant-factor (bicriteria) approximation algorithm for \textit{LUFL}.

\textbf{Theorem 2.} There is a relaxed constant-factor approximation for instances of \textit{LUFL} with uniform upper bounds (and non-uniform lower bounds).

To prove this theorem we start by presenting an LP-based bicriteria approximation for \textit{LBFL} with non-uniform lower bounds. Such approximations were known before, but ours is the first one whose cost can be compared to an LP relaxation. We emphasize that such bounds may be useful to obtain stronger results. For example, the LP-based \textit{CFL} bicriteria approximation by [1] was a key component in devising the true LP-based approximation in [4]. Perhaps our result could be used in an analogous result for \textit{LBFL}.

Next, we incorporate connectivity requirement. We obtain the first constant-factor bicriteria approximation for the \textit{connected lower-bounded facility location} problem with non-uniform lower bounds. We then extend this to a relaxed constant-factor approximation for \textit{C-LUFL} when the upper bounds \(U\) are uniform and the core cable multiplier \(M\) is \(O(U)\). Some remarks on the difficulty of extending our approach to the case \(M = \omega(U)\) are presented in the conclusion. Our second main result is the following.

\textbf{Theorem 3.} There is a relaxed constant-factor approximation for instances of \textit{C-LUFL} with uniform upper bounds where \(M = O(U)\).

A key ingredient in our approach is a clustering step to avoid the standard “filtering” steps. That is, in classic facility location and \textit{CFL} rounding algorithms a popular approach is to consider a ball around each client \(j\) whose radius is roughly the fractional cost of serving \(j\). Values \(x_{ij}\) where \(i\) lies far outside this ball are set to 0 and the remaining \(x_{ij}\) values are rounded up by a small constant factor in order to get a solution that is “concentrated” around each client. This approach fails when lower bounds are present. We develop a clustering procedure to find a set of cluster centers \(\mathcal{C}\) using a Voronoi diagram which is inspired by approaches to \textit{capacitated k-median} problem that was considered in [15, 16].

\section{LP Relaxations and Starting steps}

We present LP relaxations for \textit{LUFL} as well as \textit{C-LUFL}. For each \(i \in F\), \(y_i\) indicates if facility \(i\) is opened. For each \(i \in F\) and \(j \in D\), \(x_{ij}\) indicates if client \(j\) is assigned to facility \(i\).

\begin{align*}
\min & \sum_{i \in F} \mu_i y_i + \sum_{j \in D} \sum_{i \in F} c_{ij} x_{ij} \\
\text{(LP-LUFL)} & \sum_{i \in F} x_{ij} = 1 \quad \forall j \in D \quad (2) \\
& x_{ij} \leq y_i \quad \forall i \in F, j \in D \quad (3) \\
& \sum_{j \in D} x_{ij} \leq U_i y_i \quad \forall i \in F \quad (4) \\
& L_i y_i \leq \sum_{j \in D} x_{ij} \quad \forall i \in F \quad (5) \\
& x_{ij}, y_i \in [0, 1] \quad \forall i \in F, j \in D
\end{align*}
Constraints (2) and (3) are standard facility location constraints saying that any client has to be assigned to an open facility in an integer solution. Constraints (4) and (5) ensure the lower and upper bounds are satisfied at the open facilities.

Extending LP-LUFL to model a relaxation for C-LUFL, we let $z_e$ indicate if edge $e \in E$ is used by the core Steiner tree. We first guess one particular facility $r$ that is open in the optimum solution and we we called $r$ the root. LP-C-LUFL is an integer program formulation of C-LUFL.

$$\min \sum_{i \in F} \mu_i y_i + \sum_{j \in D} \sum_{i \in F} c_{ij} x_{ij} + M \sum_{e \in E} c_e z_e$$

(LP-C-LUFL)

\[ \sum_{e \in \delta(S)} z_e \geq \sum_{i \in S} x_{ij} \quad \forall S \subseteq V \setminus \{r\}, j \in D \]  

(6)

\[ y_r = 1 \]  

(7)

\[ x_{ij}, y_i, z_e \in [0,1] \quad \forall i \in F, j \in D, e \in E \]

Constraints (6) guarantee that all open facilities are connected to facility $r$ via core links, where Constraint (7) forces facility $r$ to be opened. Note that while (6) introduces exponentially many constraints, they can easily be separated by an efficient minimum-cut algorithm. Thus we can solve both (LP-LUFL) and (LP-C-LUFL) in polynomial time using the ellipsoid method.

2.1 Reduction Lemmas

In this section we present two lemmas that are used in the algorithms we present. The first lemma is a general clustering step that is applied as a first step of our LP rounding and reduces the facility location problem on hand to solving the problem on a specific cluster of clients facilities. This clustering step has similarities to a Voronoi diagram and for that reason we call it Voronoi clustering (inspired by [15, 16]). The second lemma shows how one can then extend the results obtained via this reduction step to the case where connectivity (with core cables) is required between open facilities.

Let $(x, y, z)$ be a feasible solution to the LP relaxation of (LP-C-LUFL). Let $L^j$ be the connection cost of client $j$ in the LP, i.e. $L^j = \sum_{i \in F} c_{ij} x_{ij}$. The general idea is to select clients in increasing order of their $L^j$ values and selecting them as centers if they are far from all centers so far. We then define a Voronoi cell with center $j$ to be the set of all facilities for which $j$ is the closest center. This Voronoi clustering will be an important tool in decomposition of an LP solution in our rounding algorithms.

The following algorithm finds a set of clients $\mathcal{C}$ that will act as the centers in the Voronoi diagram and a partition $\{P_j\}_{j \in \mathcal{C}}$ of $F$ where $i \in P_j$ means $j$ is a closest center to $i$. Here, $\lambda$ is some parameter that we can specify. Larger values mean the centers are further apart. It also records a cluster center $\delta(j) \in \mathcal{C}$ for each $j \in D$: if $j \in \mathcal{C}$ then $\delta(j) = j$ and if $j \notin \mathcal{C}$ then $\delta(j)$ is the center that caused $j$ to not be included in $\mathcal{C}$ (it may not be the closest center to $j$).
Algorithm 1. Voronoi Clustering algorithm ($\lambda$)

\[ C \leftarrow \{j^*\} \text{ where } j^* = \arg \min_j L_j; \]

for each \( j' \in D - \{j^*\} \) in increasing order of \( L_j \) do

if \( c_{ij'} > 2\lambda \cdot L_j \) for all \( j \in C \) then

\[ C \leftarrow C \cup \{j^*\}; \]

\( \delta(j') \leftarrow j'; \)

else

let \( j \in C \) be some center with \( c_{ij'} \leq 2\lambda \cdot L_j \);

\( \delta(j') \leftarrow j; \)

end

for each \( j \in C \) do

\( P_j \leftarrow \{i \in F : c_{ik} \leq c_{jk} \text{ for all } k \in C, k \neq j\}; \)

Comment: break ties arbitrarily so each \( i \in F \) lies in exactly one \( P_j \).

end

return \((C, P, \delta)\)

Note that by construction of \( \delta \) we have that \( c_{\delta(\eta)} \leq 2\lambda L_j \) for each \( j \in D \).

In Lemma 4 we show that for each center \( j \in C \), there is a facility \( i \) that is close to \( j \) whose opening cost can be paid for by the fractional opening cost paid by the LP for facilities near \( i \). Furthermore, this facility \( i \) has a small enough lower bound that we can approximately satisfy by assigning to it all fractional client demand that was sent to some facility in \( P_j \). For each client \( j \) and positive radius \( R \), we let \( B(j, R) = \{v \in V : c_{ij} \leq R\} \) be a ball centered at \( j \).

\[ \textbf{Claim 4.} \]

Let \((x, y)\) be values satisfying constraints (2), (3), and (5). Suppose \((C, P, \delta)\) is returned by calling Algorithm 1 with some given \( \lambda \). Let \( \hat{X}_j = \sum_{i \in P_j} \sum_{i' \in D} x_{ij'} \) and let \( \eta \in (1, \lambda] \). For each \( j \in C \), there exists some \( i \in B^j := B(j, \eta L_j) \) fulfilling: (i) \( \mu_i \leq \frac{2\eta}{\eta - 1} \sum_{i' \in B^j} \mu_{i'} y_{i'} \) and (ii) \( L_i \leq \frac{2\eta}{\eta - 1} \sum_{i' \in B^j} \mu_{i'} y_{i'} \).

\[ \textbf{Proof.} \]

First observe that \( \sum_{i \in B^j} y_i \geq 1 - \frac{1}{\eta} \). For each \( i \in B^j \), let \( \bar{y}_i = \frac{x_{ij}}{\sum_{i' \in B^j} x_{ij'}} \). Note that each \( x_{ij} \in B^j \),

\[ \bar{y}_i \leq \frac{\eta}{\eta - 1} x_{ij} \leq \frac{\eta}{\eta - 1} y_i, \]

holds by Constraints (3) and the fact that at least \( \frac{\eta - 1}{\eta} \) portion of \( j \)'s demand is served within \( B^j \) (using Markov’s inequality).

Now think of \( \bar{y}_i \) as a probability distribution over facilities \( B^j \) (note that \( \sum_{i \in B^j} \bar{y}_i = 1 \)). Suppose we sample a facility \( i \) from this distribution.

\[ \textbf{Claim 1.} \ Pr[\mu_i > \frac{2\eta}{\eta - 1} \sum_{i' \in B^j} \mu_{i'} y_{i'}] < 1/2. \]

\[ \textbf{Proof.} \]

Observe that \( \sum_{i' \in B^j} \mu_{i'} \bar{y}_{i'} \leq \frac{\eta}{\eta - 1} \sum_{i' \in B^j} \mu_{i'} y_{i'} \). This, with Markov’s inequality, implies: \( \Pr[\mu_i > \frac{2\eta}{\eta - 1} \sum_{i' \in B^j} \mu_{i'} y_{i'}] \leq \Pr[\mu_i > 2 \sum_{i' \in B^j} \mu_{i'} y_{i'}] < 1/2. \)

\[ \textbf{Claim 2.} \ Pr[L_i > \frac{2\eta}{\eta - 1} \bar{X}_j] < 1/2. \]

\[ \textbf{Proof.} \]

Using (5) and (8) and the fact that by choice of \( \eta \), \( B^j \subseteq P_j \) we have

\[ \bar{X}_j \geq \sum_{i \in B^j} \sum_{j' \in D} x_{ij'} \geq \sum_{i \in B^j} y_i L_i \geq \frac{\eta - 1}{\eta} \sum_{i \in B^j} L_i \bar{y}_i \]

This implies: \( \Pr[L_i > \frac{2\eta}{\eta - 1} \bar{X}_j] \leq \Pr[L_i > 2 \sum_{i \in B^j} \bar{y}_i L_i] < 1/2. \)

The above two claims immediately imply that with positive probability, there is a facility that satisfies both conditions in inequalities (i) and (ii), respectively. Hence the lemma holds. \( \square \)
Our next lemma demonstrates the utility of our clustering algorithm even in the presence of the connectivity requirements. We show below that if we find a (lower/upper bounded) facility location solution within each cluster and if we can connect those open facilities to the center of the clusters using core cables cheaply then we can connect the centers using core cables cheaply. This helps us to reduce the problem to solving each Voronoi cell separately.

Lemma 5. Let \((x, y, z)\) be values satisfying (2)-(3) and (6)-(7) and \((C, P, \delta)\) be returned by Algorithm 1 with \(x, y, z\), and some given \(\lambda\). Let \(\eta \in (1, \lambda)\). Then we can efficiently find a Steiner tree that connects \(C\) with cost at most \(\frac{\lambda}{\lambda-\eta} \cdot \frac{2\eta}{\eta-1} \cdot M \cdot \sum c_\epsilon z_\epsilon\).

The idea of the proof (which appears in Appendix) is to consider balls \(\{B(j, \eta L^2) : j \in C\}\). These balls are disjoint and one can show that looking at the cut constraints between the balls, the \(z_\epsilon\) values (from LP) satisfy cut constraints for the Steiner tree problem over these balls.

3 An LP-Based Approximation Algorithm for LUFL

In this section we present a rounding bicriteria approximation algorithm for LUFL. We start with the simpler case where we only have lower bounds and then show how to extend the algorithm to work for when there are both upper and lower bounds for facility loads.

3.1 Lower-Bounded Facility Location

We first consider the case where all facilities have infinite capacities. An LP to this case can be written as follows. We let \((x, y)\) and \(\text{OPT}_{\text{LP}}\) be the optimal solution and the optimum cost of LP-LFL, respectively.

\[
\begin{align*}
\min \sum_{i \in F} \mu_i y_i + \sum_{j \in D} \sum_{i \in F} c_{ij} x_{ij} \\
&\text{(LP-LFL)}
\end{align*}
\]

\[(2),(3),(5)\]

\[x_{ij}, y_i \geq 0\]

It is easy to see that LP-LFL has unbounded integrality gap: Consider a small instance of LBFL consisting of \(2(L-1)\) clients (with unit demands), two zero-cost facilities each collocated with \(L-1\) clients, and an edge of length \(L\) between these two facilities. While the optimal cost to IP is \(L(L-1)\), LP can manage to pay only \(2(L-1)\) by opening both facilities to the extent of \(\frac{L-1}{L}\), and thereby only sending \(\frac{1}{L}\) demand of each client to its far facility. Hence, the integrality gap can be made arbitrarily large by increasing \(L\). Therefore bicriteria approximation is unavoidable if we use LP.

Let \(\eta > 1\) be a parameter we may choose, larger values result in more expensive solutions with smaller violations in the lower bound. Our algorithm for LBFL has two steps and works as follows. We first find a Voronoi clustering using Algorithm 1 and then for each cluster center \(j\) we open one facility in the cell as guaranteed by Lemma 4. All demand \(\hat{X}_j\) that is fractionally assigned to \(P_j\) is assigned to this open facility. To turn this into an integer assignment of clients to facilities, we then compute the minimum-cost integer flow that satisfies the relaxed lower bounds. The fact that this is cheap is witnessed by the fractional assignment we find in the first part of the algorithm.
Algorithm 2: LBFL rounding

**Step 1:** Construct a Voronoi clustering $(\mathcal{C}, \mathcal{P}, \delta)$ by running Algorithm 1 with the given $x$, $y$, and $\lambda = \eta$.

**Step 2:** Let $I = \{i(j) : j \in \mathcal{C}\}$, where $i(j) \in P_j$ is the facility described in Lemma 4. Open facilities $I$ and find the cheapest assignment of clients to them such that each open facility $i$ serves at least $\frac{\eta-1}{2\eta} L_i$ demand.

▶ **Theorem 6.** Algorithm 2 computes in polynomial time a solution to LBFL with the following properties:

(i) The solution cost is at most $\max\{4(\eta + 1), \frac{2n}{\eta-1}\} \cdot \text{OPT}_{LP}$.

(ii) Each open facility $i \in I$ is serving at least $\left\lceil \frac{\eta-1}{2\eta} L_i \right\rceil$ clients.

**Proof.** We provide a solution as described in Step 2 fulfilling the claimed properties. Consider $(\mathcal{C}, \mathcal{P}, \delta)$ constructed at Step 1. Recall $X_j = \sum_{i \in P_j} \sum_{j' \in D} x_{ij'}$.

By Lemma 4 and the fact that $P_j$ cells are disjoint, the total opening cost is bounded as follows.

$$\mu(I) \leq \sum_{j \in \mathcal{C}} \mu_i(j) \leq \frac{2n}{\eta-1} \sum_{j \in \mathcal{C}} \sum_{i \in P_j} \mu_i y_i \leq \frac{2n}{\eta-1} \sum_{i \in P} \mu_i y_i.$$  \hspace{1cm} (10)

Assigning the fractional demands $X_j$ aggregated at $j$ to each $i(j) \in I$ guarantees the second property; see Lemma 4. Hence, we only need to show this assignment is cheap and how to turn it into an integer assignment of no more cost.

Consider some client $j' \in D$ and some facility $i \in F$. In what follows we show that $x_{ij'}$ units of demand travel a distance of at most $4(\eta L_j^j + c_{ij'})$. Say that $i \in P_j$ and let $B_j^j = B(j, \eta L_j^j)$. Thus, in this assignment the $x_{ij'}$-fraction of demand travels distance $c_{ij'}$.

We consider two cases:

**Case $j' \notin B_j^j$:** We have

$$c_{ij'} \leq c_{ij} + c_{ij'}$$  \hspace{1cm} (by the triangle inequality)

$$\leq \eta L_j^j + c_{ij'}$$  \hspace{1cm} (using the fact that $i(j) \in B_j^j$)

$$\leq 2c_{ij'}$$  \hspace{1cm} (using the fact that $j' \notin B_j^j$)

$$\leq 2(c_{ij} + c_{ij'})$$  \hspace{1cm} (by the triangle inequality)

$$\leq 2(c_{i\delta(j')} + c_{ij'})$$  \hspace{1cm} (using the fact that $i \in P_j$)

$$\leq 2(2\eta L_j^j + c_{ij'})$$  \hspace{1cm} (by the triangle inequality)

$$\leq 2(2\eta L_j^j + 2c_{ij'})$$  \hspace{1cm} (from the clustering procedure)

**Case $j' \in B_j^j$:** In this case $c_{ij'}(j) \leq 2\eta L_j^j$ (by the triangle inequality). Below we show that $L_j^j \leq 2L_j^j$, which immediately implies $c_{ij'}(j) \leq 4\eta L_j^j$.

▶ **Claim 3.** $L_j^j \leq 2L_j^j$.

**Proof:** Assume, for the sake of contradiction, that $L_j^j > 2L_j^j$. First observe that by the ordering clients are selected as centers in $C$, $j' \notin C$: note that $j \in C$, and since we assumed $L_j^j < L_j^j$, and because $c_{ij'} \leq 2\eta L_j^j$ (recall $j' \in B_j^j$), if $j' \in C$ then it would have prevented $j$ from being added to $C$ in the first step. Now, consider $\delta(j') \in C$. Note that $L_j^{(j')} \leq L_j^j$ and $c_{j\delta(j')} \leq 2\eta L_j^j$. This implies

$$c_{j\delta(j')} \leq c_{j\delta(j')} + c_{i\delta(j')}$$  \hspace{1cm} (by the triangle inequality)

$$\leq \eta L_j^j + 2\eta L_j^j$$  \hspace{1cm} (by $j' \in B_j^j$ and clustering procedure)

$$\leq \eta L_j^j + \eta L_j^j$$  \hspace{1cm} (using the assumption that $L_j^j > 2L_j^j$)

$$\leq 2\eta L_j^j$$
which is a contradiction because then \(\delta(j')\) would also have blocked \(j\) from being added to \(C\). The claim follows.

This completes the proof of this case that \(c'_{j'i(j)} \leq 4\eta Lj'\).

In either case, \(x_{j'}\) travels a distance of at most \(4(\eta Lj' + c_{j'})\). Thus, the total assignment cost of this fractional solution is bounded by

\[
4 \sum_{i \in F} \sum_{j' \in D} x_{j'} (\eta Lj' + c_{j'}) = 4n \sum_{j' \in D} Lj' \sum_{i \in F} x_{j'} + 4 \sum_{i \in F} \sum_{j' \in D} c_{j'} x_{j'}
\]

\[
= 4n \sum_{j' \in D} Lj' + 4 \sum_{i \in F} \sum_{j' \in D} c_{j'} x_{j'} \quad \text{using (2)}
\]

\[
= 4(\eta + 1) \sum_{i \in F} \sum_{j' \in D} x_{j'} c_{j'} \quad \text{(by def. of } Lj')
\]

Together with (10), this implies the claimed bound.

Finally, because of the integrality of flows with integer lower bounds and because we have explicitly described a cheap fractional flow from the clients to the open facilities that satisfies the integer lower bounds \(\lceil \frac{n-1}{2\eta} L_i \rceil\), then there is an integer assignment \(\sigma : D \to I\) that also satisfies these lower bounds with no greater cost. Note it could be that \(\lceil \frac{n-1}{2\eta} L_i \rceil = 0\). It is still possible to recover a an \(O(1)\)-approximation that ensures each open facility \(i \in I\) is assigned at least \(L_i/3\) clients by choosing \(\eta\) to be sufficiently large and closing any \(i \in I\) that is not assigned any clients.

For example, by choosing \(\eta = 1.28\) we get a solution of cost at most \(9.12OPT_{L,P}\) and the load of each open facility \(i\) is at least \(\lceil \frac{L_i}{3.12} \rceil\).

### 3.2 The general case with lower and upper bounds

We now consider the case where each facility has capacity \(U\) (uniform across all facilities) as well as a given lower bound \(L_i\). Let \((x, y)\) be an optimal solution to (LP-LUFL).

As before, we first use Algorithm 1 to obtain a Voronoi clustering. We then decide to open a number of facilities in each cell to route the clients demand to be served at them while satisfying the upper and lower bounds on the facility loads (approximately). The algorithm consists of two steps and works as follows.

**Algorithm 3: LUFL rounding**

**Step 1:** Construct a Voronoi clustering \((\mathcal{C}, \mathcal{P}, \delta)\) by running Algorithm 1 with the given \(x, y,\) and \(\lambda = \eta\).

**Step 2:** For each \(j \in \mathcal{C}\), we open a subset \(I_j \subseteq \mathcal{P}_j\) of facilities and send the demand served by facilities in \(\mathcal{P}_j\) (namely \(\hat{X}_j = \sum_{j'} \sum_{i \in \mathcal{P}_j} x_{ij'}\)) to those facilities as described below, depending on the value of \(\hat{X}_j\):

**Case 1.** \(\hat{X}_j \geq U\): In this case, inspired by ideas from [16], we formulate the described subproblem as another (simpler) facility location (inside the cell) using a simpler (sparse) LP.

We firstly move demand \(\hat{X}_j\) to center \(j\) as follows. For each client \(j' \in D\) and each facility \(i \in \mathcal{P}_j\), we send \(x_{ij'}\) demand from \(j'\) to \(i\) (this is what the LP is doing). Let \(\hat{d} = \sum_{j' \in D} x_{ij'}\) be the demand sent to \(i\). Next, for each facility \(i \in \mathcal{P}_j\), we send \(\hat{d}\) demand from \(i\) to \(j\). Obviously, the total cost of this moving is bounded by \(\sum_{i \in \mathcal{P}_j} \sum_{j' \in D} x_{ij'} (c_{ij'} + c_{ij})\).

We now ignore the facility lower bounds and write an LP to solve the subproblem. Solving and then rounding this LP helps us to decide which facilities in \(\mathcal{P}_j\) to open and how to assign the \(\hat{X}_j\) demand (already aggregated at \(j\)) to them. We shall show how the cost of this LP
can be bounded by the cost of the original LP restricted to this cell and an optimum solution to this LP satisfies the lower bounds on almost all facilities.

In this LP, we have a variable $\omega_i$ for each $i \in P_j$ indicating how much of the $\hat{X}_j$ is assigned to $i$.

$$\min \sum_{i \in P_j} \omega_i \left( \frac{\mu_i}{U} + c_{ij} \right)$$

$$\sum_{i \in P_j} \omega_i = \hat{X}_j$$

$$0 \leq \omega_i \leq U \quad \forall i \in P_j$$

Note that setting $\omega_i := \sum_{j \in D} x_{ij}'$ is a feasible solution with cost at most $\sum_{i \in P_j} \mu_i y_i + \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij}$ because $\sum_j x_{ij} \leq U y_i$.

Note that there is only one constraint apart from constraints $0 \leq \omega_i \leq U$. Thus, for all but one $i \in P_j$ we have $\omega_i^* \in \{0, U\}$, where $\omega^*$ indicates an optimum extreme point solution to this LP.

To round this solution $\omega^*$, let $\zeta > 1$ be a parameter we get to choose. Let $I_j = \{ i \in P_j : \omega_i^* = U \}$. If there is some $i' \in P_j$ such that $0 < \omega_i^* < U$ then add $i'$ to $I_j$ if $\omega_i^* > \frac{U}{\zeta}$. In this case, the upper bound is satisfied for every $i \in I_j$ and the lower bound is violated by no more than a $\zeta$-factor. Assign precisely $\omega_i^*$ units of demand to each $i \in I_j$. The cost of this assignment plus the cost of opening $I_j$ is at most $\sum_{i \in P_j} \zeta \mu_i y_i + \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij}$.

Otherwise, if $\omega_i^* < \frac{U}{\zeta}$ then let $i''$ be the facility in $I_j$ closest to $j$ and increase $\omega_i^*$ by $\omega_i^*$. Note that such a facility $i''$ exists because we are assuming $\hat{X}_j \geq U$. In this case, no lower bounds are violated at any $i \in I_j$ and the upper bound is violated by at most a $\left(1 + \frac{1}{\zeta}\right)$-factor. The assignment and opening cost in this case are bounded by $\sum_{i \in P_j} \mu_i y_i + \sum_{i \in P_j} \sum_{j' \in D} \frac{\zeta+1}{\zeta} x_{ij'} c_{ij}$.

In either case, we have opened $I_j$ and assigned demand to each $i \in I_j$ to satisfy the relaxed lower bounds $L_i/\zeta$ and the relaxed upper bounds $\frac{\zeta+1}{\zeta}U$. The cost of assigning $\hat{X}_j$ units of demand from $j$ to $I_j$ in this manner is at most $\zeta \sum_{i \in P_j} \mu_i y_i + \frac{\zeta+1}{\zeta} \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij}$. Altogether, the total cost (of moving the $\hat{X}_j$ demand to center $j$ plus the cost of assigning it from $j$ to facilities $I_j$) is bounded by

$$\zeta \sum_{i \in P_j} \mu_i y_i + \frac{\zeta+1}{\zeta} \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij} + \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij} \geq \zeta \sum_{i \in P_j} \mu_i y_i + \frac{2\zeta+1}{\zeta} \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij}.$$

\textbf{Lemma 7.} The total cost (over all cells of Voronoi clustering) incurred due to Case 1 of Step 2 of the algorithm is at most $\zeta \sum_{i \in F} \mu_i y_i + \frac{2\zeta+1}{\zeta} \sum_{i \in F} \sum_{j \in D} x_{ij} c_{ij}$. (See the appendix for proof.)

\textbf{Case 2.} $\hat{X}_j < U$: Observe that in this case we can simply ignore the upper bound. So (similar to that for LBFL) we open facility $i(j)$ described in Lemma 4 and send the demand to that facility as follows: For each client $j' \in D$ and each facility $i \in P_j$, we send $x_{ij'}$ demand from $j'$ (directly) to $i(j)$. Let $I_j = \{i(j)\}$ in this case. Note that facility $i(j)$ serves at least $\frac{\omega_i^*}{\zeta} L_i$.

The following bound can be obtained using the exact same arguments used to bound that in the proof of Theorem 6.

\textbf{Lemma 8.} The total cost incurred due to Case 2 of Step 2 is at most $\frac{2\zeta+1}{\zeta} \sum_{i \in F} \mu_i y_i + 4(\eta + 1) \sum_{i \in F} \sum_{j \in D} x_{ij} c_{ij}$. 

Let $I = \cup_{j \in C} I_j$ be the set of facilities opened over all Voronoi cells. Observe that each of the two cases above finds a solution to \textbf{LBFL} in which each open facility $i \in I$ serves at least \( \min \left\{ \frac{\zeta}{\eta}, \frac{\eta + 1}{2\eta} \right\} L_i \) (based on the two cases above) and at most \( \zeta + 1 \) demand.

Summing our bounds on the cost of the solutions found in each Voronoi diagram (see Lemmas 7 and 8), we see the cost of opening $I$ is at most
\[
\left( \zeta + \frac{2\eta}{\eta - 1} \right) \sum_{i \in F} \mu_i y_i, \tag{11}
\]
and the cost of assigning demands is at most:
\[
\left( \frac{2\zeta + 1}{\zeta} (2\eta + 1) + \frac{\zeta}{\zeta} + 4(\eta + 1) \right) \sum_{j \in D} \sum_{i \in F} c_{ij} x_{ij}, \tag{12}
\]
Together, (11) and (12) and using integrality of flows with integer lower and upper bounds, imply the main results of this section.

\textbf{Theorem 2 (restated).} Algorithm 3 is a polynomial time \((\rho, \alpha, \beta)\)-approximation for instances of \textbf{LUFL} with uniform capacities where \( \rho = \max \left\{ \frac{(2\zeta + 1)(2\eta + 1) + \zeta}{\zeta} + 4(\eta + 1), \frac{2\eta}{\eta - 1} + \zeta \right\}, \alpha = \max \{\zeta, \frac{2\eta}{\eta - 1}\}, \beta = \frac{\zeta + 1}{\zeta} \).

4 An LP-Based Approximation Algorithm for \textbf{C-LUFL}

In this section we show that our rounding framework for \textbf{LUFL} extends to connected variants. In the light of Lemma 5, we observe that our framework works for the connected variants too, as long as we can bound the cost of connecting facilities opened in each Voronoi cell to the center it belongs to.

We begin with the case where all facilities have infinite capacities (denoted by \textbf{C-LBFL}) and then show how to extend it to the case with lower bounds. We let \((x, y, z)\) and \(\text{OPT}_{LP}\) be the optimal solution and the optimum cost of the LP relaxation for this case, respectively.

Let \(\lambda > \eta > 1\) be constant parameters. Following the same general ideas of that for \textbf{LBFL} and using our observation described in Lemma 5, we present our algorithm for \textbf{C-LBFL} which has three stages and works as follows.

**Algorithm 4:** \textbf{C-LBFL} rounding

**Step 1:** Construct a Voronoi clustering \((C, P, \delta)\) by running Algorithm 1 with the given \(x, y, \lambda\).

**Step 2:** Open facilities \(I = \{ i(j) : j \in C \}\) and assign clients to them as described in Step 2 of Algorithm 2. Connect each facility \(i(j) \in I\) to the center it belongs to via core cables.

**Step 3:** Compute a core Steiner tree over centers \(C\) as described in Lemma 5.

One can simply adapt the proof of Lemma 5 to bound the extra cost of connecting each center \(j\) to the facility \(i(j)\) by losing a constant factor. Apart from this the proof of the following theorem is analogous to that for Theorem 6.

\textbf{Theorem 9.} Algorithm 4 computes in polynomial time a solution to \textbf{C-LBFL} with the following properties:

\(i\) The solution cost is at most \(\max \left\{ 4(\eta + 1), \frac{2\eta}{\eta - 1}, \frac{2(\lambda + \eta)\eta}{(\lambda - \eta)(\eta - 1)} \right\} \cdot \text{OPT}_{LP}\).

\(ii\) Each open facility \(i \in I\) is serving at least \(\left\lfloor \frac{\eta - 1}{2\eta} L_i \right\rfloor \) clients.

We now consider the \textbf{C-LUFL} problem. First we show that one can convert an optimum solution of \textbf{C-LUFL} to an approximate solution in which each open facility (say) \(i\) is assigned
a sufficiently large number of clients comparable not only to $U$ and $L_i$ but also to $M$ (core cable cost per unit length). This property of a near optimal solution will help use to compute approximate solutions to $C$-$LUFL$. Let $\Delta = \min\{M, U\}$. Let $OPT_{C,LU}$ be the cost of an optimal solution to $C$-$LUFL$. Observe that when the number of clients is less than $\Delta^2$, selecting only the cheapest facility to be opened and then assigning all clients to that open facility returns the optimal solution. We hence assume that the number of clients is at least $\frac{\Delta}{2}$.

**Theorem 10.** There is a feasible solution of cost at most $3OPT_{C,LU}$ to $C$-$LUFL$ in which each open facility $i$ is assigned at least $\max\{\frac{\Delta}{2}, L_i\}$ units of demand. (see the appendix for proof.)

In what follows (instead of approximating $C$-$LUFL$) we approximate the near optimal solution described above whose property is needed for our analysis to work. We write a modification of LP-$C$-$LUFL$ to model the approximate solution described above.

\[
\begin{align*}
\min & \quad \sum_{i \in F} \mu_i y_i + \sum_{j \in D} \sum_{i \in F} c_{ij} x_{ij} + M \sum_{e \in E} c_e z_e \\
\text{s.t.} & \quad \Delta y_i \leq 2 \sum_{j \in D} x_{ij} \quad \forall i \in F \quad (13) \\
& \quad x_{ij}, y_i, z_e \geq 0 
\end{align*}
\]

We let $(x, y, z)$ be the optimal solution of this LP relaxation. Let $\lambda > \eta > 1$ be constant parameters. Following the algorithm for LBFL and using Lemma 5, we extend the algorithm for C-LBFL to work for the more general case where each facility has a capacity $U$ and $M = O(U)$. Our algorithm has three steps and works as follows.

**Algorithm 5: C-LUFL rounding:**

**Step 1:** Construct a Voronoi clustering $(C, P, \delta)$ by running Algorithm 1 with the given $x$, $y$, $\lambda$.

**Step 2:** Open facilities $I = \cup_{j \in C} I_j$ and assign clients to them as described in Step 3 of Algorithm 3. Then, connect each facility $i \in I$ to the center of the cell it belongs to using core cables.

**Step 3:** Compute a core Steiner tree over centers $C$ as described in Lemma 5.

**Theorem 3 (revisited).** Algorithm 5 computes in polynomial time a $(O(1), \max\{\zeta, \frac{\Delta}{\eta - 1}\}, \frac{\Delta + 1}{\zeta})$ bicriteria approximation for instances of $C$-$LUFL$ with uniform capacities (and non-uniform lower bounds) and with $M = O(U)$.

The proof is deferred to the Appendix.

**5 Conclusion**

It would be nice to extend our approximations for C-LUFL to include the case $M = \omega(U)$. As $M$ gets larger, the cost of connecting core cables becomes so large that an optimum solution would open the fewest possible facilities, namely $k := \lceil |D|/U \rceil$. This case resembles the well-studied $k$-MST problem where it is well-known that even getting a constant-factor bicriteria approximation is not possible using the natural cut-based relaxation. So, this case poses additional difficulties.
Also open is the problem of getting constant-factor bicriteria approximations for LUFL when both lower and upper bounds are not necessarily uniform.

References

A. Missing proofs.

Proof of Lemma 5: Note that we require \( \eta < \lambda \). We assume that facility \( r \in B(j, \eta L^j) \) for some \( j \in \mathcal{C} \). The other case where \( r \notin B(j, \eta L^j) \) for any \( j \in \mathcal{C} \) is nearly identical and results in the same bound. We also observe that \( \{B(j, \eta L^j) : j \in \mathcal{C}\} \) consists of disjoint sets: if \( B(j, \eta L^j) \cap B(j', \eta L^{j'}) \neq \emptyset \) for distinct \( j, j' \in \mathcal{C} \) then \( c_{jj'} \leq 2\lambda \cdot \max\{L^j, L^{j'}\} \) so both \( j \) and \( j' \) could not be cluster centers.

Note that \( \sum_{i \in B(j, \eta L^j)} x_{ij} \geq \frac{n-1}{\eta} \) holds for any \( j \in \mathcal{C} \), using of Markov’s inequality. This, together with (6), implies that vector \( \frac{n}{\eta} z \) is a feasible fractional solution to the standard cut based LP relaxation of the Steiner tree problem with terminals being balls \( B(j, \eta L^j) \) contracted at their centers. Thus, we can efficiently find a Steiner tree \( \hat{T} \) over these contracted balls (on the resulting graph after contracting balls) with cost at most \( \frac{2n}{\eta^2} \sum c_e x_e \).

Now we have to convert this tree \( \hat{T} \) into a Steiner tree over centers \( \mathcal{C} \). When we uncontract the balls, each edge of \( \hat{T} \) between two balls around centers \( j, j' \) can be replaced with the edge between two closest nodes, say \( u \in B(j, \eta L^j) \) and \( v \in B(j', \eta L^{j'}) \). We add edges \( ju \) and \( vj' \) for each such \( uv \in \hat{T} \) to complete the Steiner tree. To bound the cost of these new edges, observe that \( \eta < \lambda \) and not only balls \( B(j, \eta L^j) \) and \( B(j', \eta L^{j'}) \) are disjoint, but also balls \( B(j, \lambda L^j) \) and \( B(j', \lambda L^{j'}) \) are disjoint as well by the same argument. So we can “charge” the cost of two new edges \( ju \) and \( vj' \) to the section of edge \( uv \) that falls between the two nested balls as follows. Let \( \alpha = \max\{L^j, L^{j'}\} \). Since \( u \in B(j, \eta L^j) \) and \( v \in B(j', \eta L^{j'}) \) then \( c_{uj} + c_{vj'} \leq 2\eta \alpha \). Furthermore, \( 2\lambda \alpha \leq c_{jj'} \leq c_{uj} + c_{uv} + c_{vj'} \leq c_{uv} + 2\eta \alpha \). Therefore,

\[
c_{uj} + c_{vj'} \leq 2\eta \alpha = \frac{2\eta}{\lambda - \eta} \cdot (\lambda - \eta) \cdot \alpha \leq \frac{\eta}{\lambda - \eta} c_{uv}.
\]

Thus, the total cost of this tree is at most \( \left( 1 + \frac{\eta}{\lambda - \eta} \right) \cdot \frac{2n}{\eta^2} M \cdot \sum c_e x_e \). \hfill □

Proof of Lemma 7: The total cost is bounded by

\[
\sum_{j \in \mathcal{C}} \left( \zeta \sum_{i \in P_j} \mu_i y_i + \frac{2\zeta + 1}{\zeta} \sum_{i \in P_j} \sum_{j' \in D} x_{ij} c_{ij} + \sum_{i \in P_j} \sum_{j' \in D} x_{ij} c_{ij'} \right)
\]

\[
= \zeta \sum_{i \in F} \mu_i y_i + \frac{2\zeta + 1}{\zeta} \sum_{j \in \mathcal{C}} \sum_{i \in P_j} \sum_{j' \in D} x_{ij} c_{ij} + \sum_{i \in P_j} \sum_{j' \in D} x_{ij} c_{ij'},
\]

(14)

using the fact that \( P_j \) cells are disjoint.

Note that for any \( i \in P_j \) and any \( j' \in D \) we have

\[
c_{ij} \leq c_{i\delta(j')} \quad \text{(using the fact that} \ i \in P_j)\]

\[
\leq c_{ij'} + c_{i\delta(j')} \quad \text{(by the triangle inequality)}\]

\[
\leq c_{ij'} + 2\eta L^j \quad \text{(from Step 1)}
\]

Hence, we have:

\[
\frac{2\zeta + 1}{\zeta} \sum_{j \in \mathcal{C}} \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} c_{ij} \leq \frac{2\zeta + 1}{\zeta} \sum_{j \in \mathcal{C}} \sum_{i \in P_j} \sum_{j' \in D} x_{ij'} (c_{ij'} + 2\eta L^j)
\]

\[
= \frac{2\zeta + 1}{\zeta} \sum_{i \in F} \sum_{j' \in D} c_{ij'} x_{ij'} + \frac{(2\zeta + 1)(2\eta)}{\zeta} \sum_{j' \in D} L^j' \quad \text{(by (2))}
\]

\[
= \frac{(2\zeta + 1)(2\eta + 1)}{\zeta} \sum_{i \in F} \sum_{j' \in D} c_{ij'} x_{ij'}.
\]

This, together with (14), implies the claimed bound. \hfill □
Proof of Theorem 10: Consider an optimum solution of C-LUFL. Let $I^*, \sigma^*$, and $T^*$ be the optimal set of open facilities, assignment function of clients, and the core network tree, respectively. In the following we construct a feasible solution to C-LUFL which costs at most three times the cost of the optimal solution, while respecting the claimed lower bounds for facilities.

We first transform $T^*$ into a TSP tour over $I^*$, call it $\Gamma$ (by doubling and short-cutting over repeated facilities). Note that $\sum_{e \in \Gamma} c_e \leq 2 \sum_{e \in T^*} c_e$. Our idea is to move around this tour and close some facilities that serve few clients and move the assigned clients to them to the next facility and show that this can be done without increasing the cost too much. Let $w_i$ be the demand that is currently assigned to facility $i \in I^*$; let $w_i = |\sigma^{-1}(i)|$ initially. For the sake of presentation, let us assume that there exists some facility $\hat{i}_0 \in I^*$ for which $w_{\hat{i}_0} = \frac{\Delta}{2} \geq L_{\hat{i}_0}$ (this can be simply enforced either by replacing some facility with two collocated copies of that facility or just by considering a chain of successive facilities of the tour as the claimed facility). Now, we traverse the tour $\Gamma$ in (say) clockwise direction and adjust the loads of the facilities one by one. Let $\hat{i}_t \in I^*$ (for $t < |I^*|$) be the $t$-th facility that is reached in the tour (after leaving $\hat{i}_0$). Starting from $\hat{i}_1$, if $w_{\hat{i}_t} < \Delta$, we send $w_{\hat{i}_t}$ units of demand from $\hat{i}_1$ to $\hat{i}_2$ via edge $\{\hat{i}_1, \hat{i}_2\} \in \Gamma$; for otherwise, we send $\max\{w_{\hat{i}_t} - U, 0\}$ units of demand from $\hat{i}_1$ to $\hat{i}_2$. We update $w_{\hat{i}_1}$ and $w_{\hat{i}_2}$ (as well as the assignment mapping), accordingly. This guarantees that $\max\{\Delta, L_{\hat{i}_1}\} \leq w_{\hat{i}_1}$ as far as $w_{\hat{i}_1} > 0$. We perform this procedure for facilities $\hat{i}_2$, $\hat{i}_3$, $\cdots$, $\hat{i}_{|I^*|-1}$ in turn, however for facility $\hat{i}_{|I^*|-1}$ we send its current load to the next facility, namely facility $\hat{i}_0$, only if $w_{\hat{i}_{|I^*|-1}} < \frac{\Delta}{2}$. After this, we simply close the facilities which are now assigned no demand. Clearly, this results a feasible solution satisfying the claimed property. It only remains then to bound the cost of this solution.

Obviously the facility opening and the core connection costs of the new solution is no more than that of the initial solution. Observe that the total demand crossing along each edge of the tour (due to the clients reassignments described above) is no more than $\Delta$. Hence, using the triangle inequalities, one can simply show that the total reassignments cost is bounded by $\Delta \cdot \sum_{e \in \Gamma} c_e + \sum_{j \in D} c_{\sigma^*(j)}j$. Note that that $\Delta \leq M$ and that $\sum_{e \in \Gamma} c_e \leq 2 \sum_{e \in T^*} c_e$. Altogether, we deduce that the total reassignments cost is bounded by $2M \sum_{e \in \Gamma} c_e + \sum_{j \in D} c_{\sigma^*(j)}j \leq 2OPT$. This concludes the proof.  

Proof (sketch) of Theorem 3:

Lemma 11. The core connection cost incurred by Step 2 of the algorithm is at most $O(1) \sum_{i \in F} \sum_{j \in D} c_{ij}x_{ij}$.

Sketch. By (13), one can simply verify that $\hat{X}_j \geq \frac{1}{2\eta} \Delta$. This implies

$$\omega_i \geq \max\{\frac{n-1}{2\eta}, \frac{1}{z}\} \Delta$$

(15)

where $w_i$ indicates the demand which is sent from center $j$ to facility $i \in I_j$ at the end of this step. (see Step 2 of Algorithm 3 for details).

The core connection cost incurred by Step 2 can be bounded as follows:

$$\sum_{j \in C} M \sum_{i \in I_j} c_{ij} \leq O(1) \sum_{j \in C} \sum_{i \in I_j} c_{ij}w_i \leq O(1) \sum_{i \in F} \sum_{j \in D} c_{ij}x_{ij}$$

using (15), the assumption that $M \leq O(1)\Delta$, and the exact same arguments used to bound the cost incurred by Step 2 of Algorithm 3.

Together, Lemma 5, Lemma 11, and the results of Theorem 2 conclude the proof.