Improved Bounds for Metric Capacitated Covering Problems

Sayan Bandyapadhyay
Department of Informatics, University of Bergen, Norway
sayan.bandyapadhyay@gmail.com

Abstract

In the Metric Capacitated Covering (MCC) problem, given a set of balls $\mathcal{B}$ in a metric space P with metric d and a capacity parameter U, the goal is to find a minimum sized subset $\mathcal{B}' \subseteq \mathcal{B}$ and an assignment of the points in P to the balls in $\mathcal{B}'$ such that each point is assigned to a ball that contains it and each ball is assigned with at most U points. MCC achieves an $O(\log |P|)$-approximation using a greedy algorithm. On the other hand, it is hard to approximate within a factor of $o(\log |P|)$ even with $\beta < 3$ factor expansion of the balls. Bandyapadhyay et al. [SoCG 2018, DCG 2019] showed that one can obtain an $O(1)$-approximation for the problem with 6.47 factor expansion of the balls. An open question left by their work is to reduce the gap between the lower bound 3 and the upper bound 6.47. In this current work, we show that it is possible to obtain an $O(1)$-approximation with only 4.24 factor expansion of the balls. We also show a similar upper bound of 5 for a more generalized version of MCC for which the best previously known bound was 9.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing → Approximation algorithms

Keywords and phrases Capacitated covering, approximation algorithms, bicriteria approximation, LP rounding

Digital Object Identifier 10.4230/LIPIcs.ESA.2020.9

Funding This work is partly supported by the Research Council of Norway via the project "MULTIVAL".

Acknowledgements I am indebted to Tanmay Inamdar for giving invaluable feedback on this work. I also thank the anonymous reviewers whose suggestions have helped to further improve the quality of the paper.

1 Introduction

In any metric space P with metric d, a ball $B(c, r)$ with center $c \in P$ and radius r is defined as the set of points at a distance at most r from c, i.e., $B(c, r) = \{p \in P \mid d(c, p) \leq r\}$. In the *Metric Capacitated Covering* (MCC) problem, we are given a set of balls $\mathcal{B}$ in the metric space P with metric d. We are also given a capacity parameter $U \in \mathbb{N}$ for the balls. The goal is to find a minimum sized subset $\mathcal{B}' \subseteq \mathcal{B}$ and an assignment $\phi : P \to \mathcal{B}'$ such that for any point $p \in P$, the ball $\phi(p)$ contains p and the number of points assigned to a ball $B \in \mathcal{B}'$ via ϕ is at most U, i.e., $|\phi^{-1}(B)| \leq U$. For $B_i \in \mathcal{B}$, we denote its center and radius by c_i and r_i, respectively.

The greedy algorithm of [28] yields an $O(\log |P|)$-approximation for MCC. Indeed, this approximation factor is tight, which can be proved using the following simple reduction from set cover. For each element, add a point. For each set, add a ball of radius 1. If an element is in a set, then the distance between the center of the corresponding ball and the corresponding point is set to 1. Consider the metric space induced by the centers and the points. The capacity of each ball is set to the total number of elements, say n. Now, if there is a set cover of size k, then all the points can be covered by k balls without violating the

28th Annual European Symposium on Algorithms (ESA 2020).
Editors: Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders; Article No. 9; pp. 9:1–9:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

capacities. The converse is also true. As set cover is hard to approximate within a factor of $o(\log n)$ under standard complexity theoretic assumptions [16], it is not possible to find an approximation for MCC which is asymptotically better than $O(\log n)$.

As it is not possible to obtain a $o(\log n)$-approximation for MCC, researchers have focused on obtaining bicriteria approximation. An (α, β) bicriteria approximation for MCC is a solution where the balls can be expanded by a factor of β (i.e., for a ball $B_i \in \mathcal{B}$ and a point p_j assigned to B_i, $d(c_i, p_j) \leq \beta \cdot r_i$) and the size of the solution is at most α times the optimum solution size (that does not expand the balls). From the above reduction, it follows that no $(o(\log n), \beta)$ bicriteria approximation is possible for MCC under standard complexity theoretic assumptions for any $\beta < 3$. This is true, as in the construction for a ball B_i that does not contain a point p_j, the distance between c_i and p_j is at least 3. Thus, with less than 3 factor expansion, B_i cannot contain any more points than before.

On the positive side, Bandyapadhyay et al. [4] obtained an $(O(1), 6.47)$ bicriteria approximation for the problem, i.e., with only a 6.47 factor expansion of the balls it is possible to obtain a constant approximation. Their algorithm is based on rounding of the natural LP relaxation of MCC. One problem that was left open by the work of [4] is to reduce the gap between the lower bound 3 and the upper bound 6.47. Thus, for what possible value of $3 \leq \beta < 6.47$ can one obtain an $(O(1), \beta)$ bicriteria approximation for MCC? They also consider a generalization of MCC – *Metric Monotonic Capacitated Covering* (MMCC). This problem is similar to MCC except each ball B_i has its individual capacity $U_i \in \mathbb{N}$ which must be satisfied if it is chosen in the solution and the capacities are *monotonic* – for any two balls B_i and B_j if the radius of B_i is at least the radius of B_j, then $U_i \geq U_j$. At first glance, this assumption might seem artificial. However, this model has applications in wireless network. In a wireless network, coverage areas of antennas can be modelled using balls. Moreover, it might be economical to invest in capacity of an antenna to serve more clients, if its coverage area is larger. Bandyapadhyay et al. [4] gave an $(O(1), 9)$ bicriteria approximation for MMCC using the same approach.

1.1 Our Results and Techniques

In this paper, we obtain improved results both for MCC and MMCC.

- For MCC, we obtain an $(O(1), 4.24)$ bicriteria approximation, i.e., it is possible to obtain an $O(1)$-approximation with only 4.24 factor expansion of the balls when the capacities are uniform.
- For MMCC, we obtain an $(O(1), 5)$ bicriteria approximation, i.e., it is possible to obtain an $O(1)$-approximation with only 5 factor expansion of the balls when the capacities are monotonic.

Similar to [4] our results are also based on LP rounding. Indeed, our starting point is their rounding algorithm. For the purpose of giving an overview of our technique, let us focus on MMCC. The algorithm in [4] consists of three steps – Preprocessing, Cluster Formation and Selection of Balls. Each of Preprocessing and Selection of Balls incurs an overhead of a factor 3 expansion of the balls, resulting in the 9 factor expansion. In our algorithm we judiciously avoid the preprocessing step to save the factor 3 expansion. At first glance, it is not entirely clear how to do the rounding without preprocessing, as the preprocessed solution has several "nice" properties. Nevertheless, we partition the set of points into two subsets and construct two auxilliary LPs. Using the initial fractional LP solution, we construct two feasible fractional solutions to these two LPs. We round these two solutions independently to obtain two integral solutions corresponding to the two subsets of points. For rounding the first LP, we use an algorithm similar to the one in [4], but without preprocessing. We show that the constructed fractional LP solution has equally nice properties so that the algorithm in [4] can be extended in this case. For rounding the second LP, we use a rather simple algorithm.

The sets of balls involved in two LPs are not necessarily disjoint, and thus a ball can be selected in both of the solutions. But, taking multiple copies of a ball is not allowed. To resolve this issue, we first identify a subset of balls and allow only these balls to be involved in both solutions. Moreover, we scale down the capacities of these balls by a suitable factor. This ensures that even if a ball is selected in both solutions, the total capacity used by the copies does not exceed the original capacity. Note that the scaling of capacities leads to a new issue that the capacities no longer satisfy the monotonicity property in general. However, we show that it is possible to overcome this issue by considering two classes of balls separately – one whose capacities remain unchanged and the other whose capacities are scaled down.

1.2 Related Work

Considering the hardness of MCC, researchers have studied the Euclidean version of the problem with the goal of obtaining better approximation. The dimension d of the space is assumed to be a constant. One interesting case is when the set $\mathcal{B}$ contains all possible unit balls, which appeared in the Sloan Digital Sky Survey project [25]. Ghasemi and Razzazi [18] have obtained a PTAS for this case. In the general Euclidean case the best known approximation factor is still $O(\log n)$. Bandyapadhyay et al. [4] showed that in this special case of MCC only $1+\epsilon$ expansion of the balls is sufficient to obtain a constant approximation.

MCC is a special version of *Capacitated Set Cover* (CSC). CSC is similar to set cover except each set S_i has a capacity U_i. Moreover, we want to find an assignment of the points to the chosen subfamily of sets such that each element is assigned to a set it is in and at most U_i elements are assigned to each set S_i. CSC is a well-studied problem. Wolsey [28] designed a greedy algorithm for CSC that achieves a tight $O(\log n)$-approximation. Capacitated vertex cover is another special case of CSC, where each element is contained in exactly two sets. A 3-approximation for this problem was given by Chuzhoy and Naor [12]. The approximation factor was subsequently improved to 2 by Gandhi et al. [17]. The generalization where each element belongs to at most a bounded number of sets is also well-studied [20, 29].

The uncapacitated version of MCC (*Metric Uncapacitated Covering* (MUC)), where each set can be assigned with any number of points is another extensively studied problem. Note that the same bicriteria hardness of MCC mentioned above holds even for MUC. But, using a simple LP rounding scheme one can obtain a $(1, 3)$ bicriteria approximation for this problem. The MUC problem in the fixed-dimensional Euclidean space also has received huge attention from the researchers. Brönnimann and Goodrich [7] have designed an $O(1)$-approximation for this problem in the plane. In a celebrated work, Mustafa and Ray [26] improved this result by obtaining a PTAS for the problem. In dimension more than 2, the problem is notoriously hard and the best known approximation is $O(\log n)$. Considering this situation Har-Peled and Lee [19] gave a $(1+\epsilon, 1+\epsilon)$ bicriteria approximation.

Capacitated clustering and facility location problems are another set of interesting and well-studied problems. One such interesting problem is capacitated k-center. $O(1)$-approximations are known both for the uniform [6, 21] and non-uniform [2, 14] version of this problem. Another popular clustering problem is capacitated k-median for which no $O(1)$-approximation is known so far. Seemingly the existing techniques are not capable of handling the combination of the global constraint on the number of centers and the capacity constraint. Indeed, if either of these constraints is allowed to be violated by an

$O(1)$ factor, $O(1)$-approximations are known in those cases [9, 8, 10, 13, 15, 23, 24]. For capacitated facility location $O(1)$-approximations are known based on local search paradigm [1, 5, 11, 22, 27] and rounding of LP [3].

1.3 Paper Outline

In Section 2 we describe the natural LP for MMCC and have some definitions, which will be useful throughout the paper. In Section 3 we give an overview of the algorithm of [4]. Our LP rounding algorithm for MMCC and the analysis appear in Section 4. In Section 5 we show how to modify our algorithm for MMCC in the uniform case to obtain the improved bound. Finally, in Section 6 we conclude with some open problems.

2 Preliminaries

Recall that in MMCC we are given a set of points P and a set of balls $\mathcal{B}$. The capacity of each ball $B_i \in \mathcal{B}$ is U_i. Also, these capacities satisfy monotonicity, i.e., for any two balls B_i and B_j, if $r_i \geq r_j$, $U_i \geq U_j$.

The relaxation of the natural LP for MMCC is shown in the following. In the LP for MMCC, we have a variable y_i for each ball $B_i \in \mathcal{B}$ that indicates whether B_i is in the solution ($y_i = 1$) or not ($y_i = 0$). For each ball B_i and each point $p_j \in P$, there is a variable x_{ij} that indicates whether p_j is assigned to B_i ($x_{ij} = 1$) or not ($x_{ij} = 0$). Constraint 1 ensures that if a point is assigned to a ball, the ball must be selected in the solution. Constraint 2 ensures that the total number of points assigned to B_i is at most U_i. Constraint 3 ensures that each point is assigned to exactly one ball. Constraint 4 ensures that if a point p_j is assigned to a ball B_i, p_j must be contained in B_i. The remaining constraints are relaxed in MMCC-LP, which define the domains of the variables. We note that the LP relaxation for MCC is same as MMCC-LP except there all the U_i are equal.

$$\text{minimize} \quad \sum_{B_i \in \mathcal{B}} y_i \qquad \text{(MMCC-LP)}$$

$$\text{s.t.} \quad x_{ij} \leq y_i \qquad \forall p_j \in P,\ \forall B_i \in \mathcal{B} \tag{1}$$

$$\sum_{p_j \in P} x_{ij} \leq y_i \cdot U_i \qquad \forall B_i \in \mathcal{B} \tag{2}$$

$$\sum_{B_i \in \mathcal{B}} x_{ij} = 1 \qquad \forall p_j \in P \tag{3}$$

$$x_{ij} = 0 \qquad \forall p_j \in P,\ \forall B_i \in \mathcal{B} \text{ such that } p_j \notin B_i \tag{4}$$

$$x_{ij} \geq 0 \qquad \forall p_j \in P,\ \forall B_i \in \mathcal{B} \tag{5}$$

$$0 \leq y_i \leq 1 \qquad \forall B_i \in \mathcal{B} \tag{6}$$

We denote any solution to MMCC-LP by (x, y). To distinguish between two different solutions, we use different annotations with x and y. The cost of (x, y) is defined as, $\text{cost}(x, y) = \sum_{B_i \in \mathcal{B}} y_i$. For an integral solution, the cost is exactly the number of balls in the solution. Consider any solution (x, y) to MMCC-LP. For a ball B_i and a point p_j, if $x_{ij} > 0$, we say B_i *serves* p_j and p_j receives x_{ij} amount of *flow* from B_i. The *flow out* of B_i is the total amount of flow $\sum_{p_j \in P} x_{ij}$ that B_i gives to all the points. Next, we define an operation that we call "reroute". For a point p_j and two balls B_i and B_ℓ, rerouting of f amount of flow for p_j from B_i to B_ℓ means we increase $x_{\ell j}$ by f and decrease x_{ij} by f. For two balls B_i and B_ℓ, rerouting of flow from B_i to B_ℓ means for each point p_j served by B_i, we reroute x_{ij} amount of flow for p_j from B_i to B_ℓ. Thus, the flow out of B_i becomes 0 after this operation. For a point p_j, a set of balls S and a ball $B_\ell \notin S$, rerouting of f amount of flow from the balls in S to B_ℓ means we increase $x_{\ell j}$ by f and decrease x_{ij} by $f_i \geq 0$ for each $B_i \in S$ such that $\sum_{B_i \in S} f_i = f$.

3 Overview of the Algorithm of [4]

Our algorithm is based on the algorithm of [4]. In this section we give an overview of the algorithm of [4]. Let (x, y) be a feasible solution to MMCC-LP. The LP rounding algorithm of [4] rounds the solution so that y values of all the balls become integral. We note that it is sufficient to obtain such a solution. Indeed, as all the capacities are integral, it is possible to find another solution with the same y values where all the x values are also integral [12]. The algorithm has three major steps. The first step is the preprocessing step. Fix a $0 < \alpha \leq 3/8$. A ball B_i is called *heavy* if $y_i = 1$ and *light* if $0 \leq y_i \leq \alpha$. Let $\mathcal{H}$ and $\mathcal{L}$ be the respective set of heavy and light balls. We note that the sets of heavy and light balls are always defined w.r.t. an LP solution. But, for simplicity we do not explicitly mention that in the notations $\mathcal{H}$ and $\mathcal{L}$. The implicit solution w.r.t. which $\mathcal{H}$ and $\mathcal{L}$ are defined can be easily derived from the context. Now, it might not be true that for all $p_j \in P$, the sum of the y values of the balls in $\mathcal{L}$ that serve p_j is at most α. In the preprocessing step, the algorithm of [4] modifies the computed LP solution to obtain another LP solution such that the above mentioned property is satisfied. In particular, they prove the following lemma.

▶ **Lemma 1** (Lemma 3.1 of [4]). *Given a feasible LP solution $\sigma = (x, y)$, and a parameter $0 < \alpha \leq \frac{3}{8}$, there exists a polynomial time algorithm to obtain another LP solution $\overline{\sigma} = (\overline{x}, \overline{y})$ that satisfies all the constraints of MMCC-LP (Constraints 1-6), except Constraint 4. Additionally, $\overline{\sigma}$ satisfies the following properties.*

1. *Any ball $B_i \in \mathcal{B}$ with non-zero $\overline{y_i}$ is either heavy ($\overline{y_i} = 1$) or light ($0 < \overline{y_i} \leq \alpha$).*
2. *For each point $p_j \in P$, we have that*
$$\sum_{B_i \in \mathcal{L} : \overline{x}_{ij} > 0} \overline{y_i} \leq \alpha, \tag{7}$$
where $\mathcal{L}$ is the set of light balls with respect to $\overline{\sigma}$.
3. *For any heavy ball B_i, and any point $p_j \in P$ served by B_i, $d(c_i, p_j) \leq 3r_i$.*
4. *For any light ball B_i, and any point $p_j \in P$ served by B_i, $d(c_i, p_j) \leq r_i$.*
5. $\text{cost}(\overline{\sigma}) \leq \frac{1}{\alpha}\text{cost}(\sigma)$.

Note that a point p_j can be fractionally assigned by the algorithm in Lemma 1 to a heavy ball B_i even if $p_i \notin B_i$, but, in this case $d(c_i, p_j)$ must be at most $3r_i$. Hence, a factor 3 expansion of the ball is sufficient for it to serve the point. In summary, the preprocessing step implicitly incurs an expansion factor of 3 for the heavy balls with respect to the new LP solution $\overline{\sigma}$. We also note that the preprocessing algorithm uses the fact that the capacities are monotonic.

The second step of the algorithm is the key step and is called Cluster Formation. In the following, we give an overview of this step. The algorithm maintains an LP solution $\overline{\sigma} = (\overline{x}, \overline{y})$ which is initially the output of the preprocessing step. This solution is essentially altered throughout the step and when the step finishes $\overline{y}_i \in \{0, 1\}$ for all $B_i \in \mathcal{B}$. Each heavy ball B_i forms a cluster which initially consists of itself ($\{B_i\}$). For any light ball B_t, either B_t is opened fully in the solution or it joins a cluster of a heavy ball by rerouting its flow to the heavy ball. The algorithm runs for several iterations until the fate of all these light balls are decided.

In each iteration, every heavy ball uses its available capacity to reroute the flow of as many intersecting light balls as possible to itself. Each such light ball joins the cluster of the heavy ball. From the remaining light balls whose fate are not yet decided, a ball is selected greedily to be included in the solution. Also, for points inside the selected ball, an appropriate amount of flow is rerouted from other balls to this ball to utilize its capacity. We skip the details of this flow rerouting in this overview. This completes the overview of the step.

Note that the flow rerouting from heavy balls to a light ball when the light ball is opened fully, is an essential component of the analysis for obtaining the constant factor guarantee on the size of the solution. Consider a light ball B_t which is selected for opening fully and assume that it serves $k_t \leq U_t$ points. Then, we can set the $\overline{x}_{tj}$ value for each of these k_t points to 1, i.e., we fully assign p_j to B_t. Note that preprocessing ensures that $\sum_{B_i \in \mathcal{L}: \overline{x}_{ij}>0} \overline{y_i} \leq \alpha$ or $\sum_{B_i \in \mathcal{H}: \overline{x}_{ij}>0} \overline{y_i} \geq 1-\alpha$. Thus, when these points are fully assigned to B_t, at least $(1-\alpha)k_t$ amount of flow is rerouted from the heavy balls to B_t which they can now use to reroute flow from other light balls. This argument is essential in the analysis. Now, we have an observation which follows due to the way light balls are added to a cluster.

▶ **Observation 2.** *Consider a cluster of a heavy ball B_h that contains the light balls $B_1, \ldots, B_\ell$. Then, when the Cluster Formation finishes,*

1. *For each $1 \leq i \leq \ell$, there is a point p_j such that $p_j \in B_h \cap B_i$.*
2. $\sum_{i=1}^{\ell} \sum_{j \in P} \overline{x}_{ij} \leq U_h - \sum_{j \in P} \overline{x}_{hj}$*, i.e., the total amount of flow out of the balls in the cluster of B_h is at most U_h.*

The third step is called Selection of Balls. In this step, from each cluster a ball is carefully selected and expanded so that it can serve all the points served by the balls in the cluster. For a cluster of a heavy ball B_h, if it is the largest ball in the cluster then B_h is selected and with three factor expansion it can serve all the points served by the cluster. As during preprocessing the heavy ball might have been expanded by a factor of 3, its total expansion factor is 9. If B_h is not the largest ball, the largest ball B_ℓ is a light ball of the cluster. Then, we select this light ball and expand by a factor of 5 so that it can serve all the points served by the cluster. The light ball can serve the total flow assigned to the cluster, as $U_\ell \geq U_h$ due to monotonicity. This is another place where the monotonicity assumption on the capacities is necessary.

The following lemma that states the guarantee achieved by the above algorithm follows due to the analysis of [4].

▶ **Lemma 3.** *There is a $(6+5\alpha)/\alpha$-approximation for MMCC that expands the balls by at most a factor of* 9*.*

4 The Modified Algorithm for MMCC

In this section, we describe our algorithm. Note that among the 9 factor expansion needed in the algorithm of [4] 3 factor is contributed by the preprocessing step. Our algorithm avoids this preprocessing step to save this factor 3 expansion.

Fix $0 < \alpha \leq 1/60$. We first compute a fractional LP solution $\sigma^* = (x^*, y^*)$ to MMCC-LP. Set $y_i = 1$ if $y_i^* > \alpha$, otherwise $y_i = y_i^*$. Also, set $x = x^*$. Note that $\sigma = (x, y)$ is a feasible solution to MMCC-LP such that $\text{cost}(\sigma) \leq \text{cost}(\sigma^*)/\alpha$. We define the sets $\mathcal{H}$ and $\mathcal{L}$ of heavy and light balls w.r.t. σ in the same way, i.e., $\mathcal{H} = \{B_i \mid y_i = 1\}$ and $\mathcal{L} = \{B_i \mid 0 < y_i \leq \alpha\}$. Note that in σ, any ball that gives some flow to a point is either a heavy or a light ball. We take one copy of the set of heavy balls and two copies of the set of light balls. Let these sets be $\mathcal{H}_1$, $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively.

Next, we partition the point set into two subsets. Let P_1 be the subset of points such that $p_j \in P_1$ if $\sum_{B_i \in \mathcal{L}} x_{ij} \leq 4\alpha$, i.e., p_j gets a flow of at most 4α from the balls in $\mathcal{L}$. Let $P_2 = P \setminus P_1$. Based on these sets P_1, P_2, we are going to construct two LP solutions to two auxilliary LPs and round them independently. Finally, we combine these two solutions to find a solution to MMCC-LP where for each $B_i \in \mathcal{B}$, $y_i \in \{0, 1\}$. Intuitively, we satisfy the demands of these two sets of points independently. The light balls are involved in both of the solutions and they might get opened fully in both of the solutions. However, we are not allowed to open multiple copies of a ball. To avoid this situation we reduce the capacity of the light balls by appropriate factor in the auxilliary LP.

Let the new capacity $U'_i = U_i/10$ for each light ball B_i. The new capacity of each heavy ball B_i remains same as before, i.e., $U'_i = U_i$. At this point the reader might wonder about the value of the scaling factor. We note that it is carefully chosen through back calculation to ensure that the analysis goes through. The first auxilliary LP that we consider is as follows.

$$\begin{aligned}
\text{minimize} \quad & \sum_{B_i \in \mathcal{L}_1 \cup \mathcal{H}_1} y_i & & \text{(AUX-LP1)} \\
\text{s.t.} \quad & x_{ij} \leq y_i & \forall p_j \in P_1,\ \forall B_i \in \mathcal{L}_1 \cup \mathcal{H}_1 & \quad (8) \\
& \sum_{p_j \in P_1} x_{ij} \leq y_i \cdot U'_i & \forall B_i \in \mathcal{L}_1 \cup \mathcal{H}_1 & \quad (9) \\
& \sum_{B_i \in \mathcal{L}_1 \cup \mathcal{H}_1} x_{ij} = 1 & \forall p_j \in P_1 & \quad (10) \\
& x_{ij} = 0 & \forall p_j \in P_1,\ \forall B_i \in \mathcal{L}_1 \cup \mathcal{H}_1 \text{ such that } p_j \notin B_i & \quad (11) \\
& x_{ij} \geq 0 & \forall p_j \in P_1,\ \forall B_i \in \mathcal{L}_1 \cup \mathcal{H}_1 & \quad (12) \\
& 0 \leq y_i \leq 1 & \forall B_i \in \mathcal{L}_1 \cup \mathcal{H}_1 & \quad (13)
\end{aligned}$$

Note that the above LP has a variable y_i for each ball B_i in $\mathcal{L}_1 \cup \mathcal{H}_1$, and a variable x_{ij} for each ball B_i in $\mathcal{L}_1 \cup \mathcal{H}_1$ and each point $p_j \in P_1$. We are not going to solve this LP. Instead, we construct a solution to this LP using σ and round it using an algorithm similar to the one in [4]. This LP is used to compare the cost of the rounded solution and the cost of σ^* in the end.

We construct an LP solution $\overline{\sigma} = (\overline{x}, \overline{y})$ from σ in the following manner. For $B_i \in \mathcal{H}_1$, $\overline{y}_i = y_i$. For $B_i \in \mathcal{L}_1$, $\overline{y}_i = 10 \cdot y_i \leq 10\alpha < 1$ ($\alpha \leq 1/60$). For $p_j \in P_1$, $B_i \in \mathcal{L}_1 \cup \mathcal{H}_1$, $\overline{x}_{ij} = x_{ij}$.

▶ **Lemma 4.** *$\overline{\sigma} = (\overline{x}, \overline{y})$ is a feasible solution to AUX-LP1 with cost at most* $\text{cost}(\sigma^*)/\alpha$.

Proof. First note that,

$$\text{cost}(\overline{\sigma}) = \sum_{B_i \in \mathcal{H}_1} y_i + 10 \sum_{B_i \in \mathcal{L}_1} y_i \leq (1/\alpha) \sum_{B_i \in \mathcal{H}_1} y_i^* + 10 \sum_{B_i \in \mathcal{L}_1} y_i^* \leq \text{cost}(\sigma^*)/\alpha.$$

For $p_j \in P_1$, $B_i \in \mathcal{L}_1 \cup \mathcal{H}_1$, $\overline{x}_{ij} = x_{ij} \leq y_i \leq \overline{y}_i$. Thus, Constraint 8 is satisfied.

For $B_i \in \mathcal{H}_1$, $\sum_{p_j \in P_1} \overline{x}_{ij} = \sum_{p_j \in P_1} x_{ij} \leq y_i \cdot U_i = \overline{y}_i \cdot U'_i$. For $B_i \in \mathcal{L}_1$, $\sum_{p_j \in P_1} \overline{x}_{ij} = \sum_{p_j \in P_1} x_{ij} \leq y_i \cdot U_i = (10 \cdot y_i) \cdot (U_i/10) = \overline{y}_i \cdot U'_i$. Thus, Constraint 9 is satisfied.

For $p_j \in P_1$, $\sum_{B_i \in \mathcal{L}_1 \cup \mathcal{H}_1} \overline{x}_{ij} = \sum_{B_i \in \mathcal{L}_1 \cup \mathcal{H}_1} x_{ij} = 1$. Thus, Constraint 10 is satisfied. Also, it is trivial to verify that Constraints 11-13 are also satisfied. Hence, the lemma follows. ◀

Next, we describe our second auxilliary LP. Let us again consider the solution $\sigma = (x, y)$ to MMCC-LP and the set of light balls $\mathcal{L}$ w.r.t. σ. Also, consider the second copy $\mathcal{L}_2$ of the set of light balls. For each point p_j in P_2, define the demand $d_j = \sum_{B_i \in \mathcal{L}_2} x_{ij}$.

$$
\begin{aligned}
\text{minimize} \quad & \sum_{B_i \in \mathcal{L}_2} y_i && && \text{(AUX-LP2)}\\
\text{s.t.} \quad & x_{ij} \le y_i && \forall p_j \in P_2,\ \forall B_i \in \mathcal{L}_2 && (14)\\
& \sum_{p_j \in P_2} x_{ij} \le y_i \cdot U_i' && \forall B_i \in \mathcal{L}_2 && (15)\\
& \sum_{B_i \in \mathcal{L}_2} x_{ij} \ge d_j && \forall p_j \in P_2 && (16)\\
& x_{ij} = 0 && \forall p_j \in P_2,\ \forall B_i \in \mathcal{L}_2 \text{ such that } p_j \notin B_i && (17)\\
& x_{ij} \ge 0 && \forall p_j \in P_2,\ \forall B_i \in \mathcal{L}_2 && (18)\\
& 0 \le y_i \le 1 && \forall B_i \in \mathcal{L}_2 && (19)
\end{aligned}
$$

Note that the above LP has a variable y_i for each ball B_i in $\mathcal{L}_2$ and a variable x_{ij} for each ball B_i in $\mathcal{L}_2$ and each point $p_j \in P_2$. Again we are not going to solve this LP. Instead, we construct a solution to this LP using σ and round it. This LP is used to compare the cost of the rounded solution and the cost of σ^* in the end.

We construct an LP solution $\hat{\sigma} = (\hat{x}, \hat{y})$ from σ in the following manner. For $B_i \in \mathcal{L}_2$, $\hat{y}_i = 10 \cdot y_i \le 10\alpha < 1$. For $p_j \in P_2$, $B_i \in \mathcal{L}_2$, $\hat{x}_{ij} = x_{ij}$.

▶ **Lemma 5.** *$\hat{\sigma} = (\hat{x}, \hat{y})$ is a feasible solution to AUX-LP2 with cost at most* $10 \cdot \text{cost}(\sigma^*)$.

Proof. First note that $\text{cost}(\hat{\sigma}) \le 10 \sum_{B_i \in \mathcal{L}_2} y_i = 10 \sum_{B_i \in \mathcal{L}_2} y_i^* \le 10 \cdot \text{cost}(\sigma^*)$. For $p_j \in P_2$, $B_i \in \mathcal{L}_2$, $\hat{x}_{ij} = x_{ij} \le y_i < \hat{y}_i$. Thus, Constraint 14 is satisfied.

For $B_i \in \mathcal{L}_2$, $\sum_{p_j \in P_2} \hat{x}_{ij} = \sum_{p_j \in P_2} x_{ij} \le y_i \cdot U_i = (10 \cdot y_i) \cdot (U_i/10) = \hat{y}_i \cdot U_i'$. Thus, Constraint 15 is satisfied.

For $p_j \in P_2$, $\sum_{B_i \in \mathcal{L}_2} \hat{x}_{ij} = \sum_{B_i \in \mathcal{L}_2} x_{ij} = d_j$. Thus, Constraint 16 is satisfied. Also, it is trivial to verify that Constraints 17-19 are also satisfied. Hence, the lemma follows. ◀

In the following, we give two algorithms for rounding the two auxilliary LPs. The rounded solution of the first LP satisfies all the constraints except the coverage constraint. The rounded solution of the second LP satisfies all the constraints except the coverage and capacity constraints. Then, we merge these two solutions to obtain a solution for MMCC-LP that does not violate any capacity constraints.

4.1 Rounding the First Auxilliary LP

Note that we are given a feasible LP solution $\overline{\sigma} = (\overline{x}, \overline{y})$ to AUX-LP1 that has the following properties.

1. For any $B_i \in \mathcal{H}_1$, $\overline{y}_i = 1$.
2. For any $B_i \in \mathcal{L}_1$, $\overline{y}_i \le 10\alpha$.
3. For any $p_j \in P_1$, $\sum_{B_i \in \mathcal{H}_1} \overline{x}_{ij} \ge 1 - 4\alpha$.
4. $\text{cost}(\overline{\sigma}) \le \text{cost}(\sigma^*)/\alpha$.

Note that Property (3) above states that for any point $p_j \in P_1$, the flow received by p_j from the balls in $\mathcal{H}_1$ is at least $1 - 4\alpha$. We will heavily use this property while performing the rounding. Indeed, we are going to use an algorithm similar to the one in [4] without the preprocessing step. In the algorithm of [4], preprocessing ensures that for any point p_j, the sum of the y values of the light balls that give non-zero flow to p_j is at most α. Note that this might not be true in our case for balls in $\mathcal{L}_1$. At first glance it is not clear how to do the rounding without this assumption. However, as we show, a similar rounding scheme can be designed using the weaker assumption on the flow mentioned above. Another hurdle to adapt the algorithm of [4] is the monotonicity assumption, which might not be true in our case because of scaling of the capacities. However, we note that only light balls' capacities are scaled by a uniform constant scaling factor. Due to this fact, we show that their algorithm can be modified to handle our case. Next, we describe our rounding algorithm.

The first step in our algorithm is Cluster Formation. In this step, for each ball $B_i \in \mathcal{L}_1$, either B_i is opened fully (added to a set $\mathcal{O}$) and flow from other balls including the balls in $\mathcal{H}_1$ are rerouted to B_i only for points in B_i. Otherwise, B_i joins a cluster of a ball in $\mathcal{H}_1$ to which its entire flow is rerouted. $\mathcal{O}$ is initialized to the empty set. For each ball $B_i \in \mathcal{H}_1$, initialize the cluster of B_i, cluster(B_i) to $\{B_i\}$. During the course of the algorithm, let $\Lambda \subseteq \mathcal{L}_1$ be the set of balls which are not yet added to $\mathcal{O}$ or to a cluster of a ball in $\mathcal{H}_1$. Throughout the algorithm, we maintain the invariant that for any point p_j which is served by a ball in Λ, p_j receives a flow of at least $1 - 4\alpha$ from the balls in $\mathcal{H}_1$. Note that in the beginning of the algorithm this is true, as $\Lambda = \mathcal{L}_1$. At any point, the available capacity of a ball B_i, $AC(B_i) = U'_i - \sum_{j \in P_1} \overline{x}_{ij}$. While the set Λ is non-empty, apply the following steps.

While there is a ball $B_i \in \mathcal{H}_1$ and $B_{i'} \in \Lambda$ such that B_i intersects $B_{i'}$ and $AC(B_i)$ is at least the flow out $\sum_{j \in P_1} \overline{x}_{i'j}$ of $B_{i'}$, reroute the flow from $B_{i'}$ to B_i. Add $B_{i'}$ to cluster(B_i). If Λ becomes empty at this point, go to the Selection of Balls stage.

For any ball $B_j \in \Lambda$, let $\mathcal{A}_j$ be the set of points currently being served by B_j. Also, let $k_j = \min\{U'_j, |\mathcal{A}_j|\}$. We add a ball $B_t \in \Lambda$ to $\mathcal{O}$ such that k_t is the maximum over all k_j for $B_j \in \Lambda$.

Next we assign points up to larger extents to B_t to utilize its capacity. There are three cases.

1. $k_t > 2$. Note that the flow out of B_t, $\sum_{j \in P_1} \overline{x}_{tj} \le 10\alpha U'_t$. Also, as $\overline{x}_{tj} = x_{tj} \le y_t \le \alpha$, $\sum_{j \in P_1} \overline{x}_{tj} \le \alpha|\mathcal{A}_t| \le 10\alpha|\mathcal{A}_t|$. Thus, $AC(B_t) \ge (1-10\alpha)k_t$. In this case, we arbitrarily select $\lfloor (1-10\alpha)k_t \rfloor$ points served by B_t and for each of those points p_ℓ, we reroute the maximum (whole) amount of flow possible from all other balls to B_t. Note that p_ℓ is no longer served by a ball in Λ, and thus the invariant is satisfied.
2. $1 \le k_t \le 2$. If $U'_t \ge |\mathcal{A}_t|$, then $|\mathcal{A}_t| = k_t$. In this case, for each of the k_t points served by B_t, we reroute the maximum amount of flow possible from all other balls to B_t. In the other case, $U'_t < |\mathcal{A}_t|$. Now, $AC(B_t) \ge (1-10\alpha)U'_t \ge 1 - 10\alpha$. The last inequality follows, as $U'_t \ge 1$. We arbitrarily select a point p_ℓ that is being served by B_t and reroute its flow from Λ to B_t. Let f be the amount of flow that now p_ℓ receives from B_t. Note that $f \le 4\alpha$. Also, p_ℓ is no longer served by a ball in Λ. Now, $AC(B_t) \ge 1 - 10\alpha - 4\alpha = 1 - 14\alpha$. We reroute $\min\{AC(B_t), 1-f\}$ amount of flow from $\mathcal{H}_1$ to B_t for p_ℓ. In any case, the points whose flow are routed to B_t in this step are no longer served by a ball in Λ, and thus the invariant is satisfied.
3. $0 < k_t < 1$. Note that, as $|\mathcal{A}_t| \ge 1$, $k_t = U'_t < 1$. Now, $AC(B_t) \ge (1-10\alpha)U'_t$. Consider any arbitrary point p_ℓ that is being served by B_t. First, reroute its flow from Λ to B_t. $AC(B_t) \ge (1-10\alpha)U'_t - 4\alpha$. Note that after this rerouting, p_ℓ is no longer served by balls in Λ, and thus the invariant is satisfied. Let p_ℓ gets a flow of f_1 from the balls in $\mathcal{H}_1$. By the invariant we maintain, f_1 is at least $1-4\alpha$. Reroute $\min\{AC(B_t), f_1\}$ amount of flow of p_ℓ from the balls in $\mathcal{H}_1$ to B_t.

When the while loop terminates each ball in $\mathcal{L}_1$ is either in $\mathcal{O}$ or added to a cluster. For each $B_i \in \mathcal{O}$, we set $\overline{y}_i = 1$ and cluster$(B_i) = \{B_i\}$.

We note that the third case ($0 < k_t < 1$) mentioned above does not occur in the context of [4], as in their case for each ball B_j, both U_j and $|\mathcal{A}_j|$ are at least 1. This case appears to be the bottleneck for our algorithm and leads to a larger constant of approximation as we will describe in the analysis.

The Selection of Balls step is more interesting in our case as the monotonicity property no longer holds in general. For a cluster of a ball in $\mathcal{O}$, we trivially select this ball. Consider the cluster of any ball $B_h \in \mathcal{H}_1$. If B_h is one of the top 10 largest balls in the cluster, then select all the balls larger than B_h and also B_h. Only B_h is expanded by a factor of 3. The flow rerouted from any selected ball of $\mathcal{L}_1$ to B_h in the Cluster Formation step is assigned to it. Note that for the remaining balls of $\mathcal{L}_1$ which are in the same cluster and not chosen, are smaller than B_h, and thus can be covered by a factor 3 expansion of B_h. The remaining flow is assigned to B_h. Otherwise, the top 10 largest balls are selected all of which are in $\mathcal{L}_1$. The flow rerouted from any selected ball to B_h in the Cluster Formation step is assigned to the ball. Now consider the remaining flow assigned to the cluster. Also consider a point p_j which receives a part of this flow and not in any of the selected balls. Then, by 5 factor expansion, any selected ball can cover p_j. We expand each selected ball by 5 factor and the remaining flow is assigned arbitrarily to selected balls respecting their capacity.

4.1.1 Analysis

Let I be the number of iterations of the outermost while loop. Also, let L_t be the ball of $\mathcal{L}_1$ added to $\mathcal{O}$ at iteration $1 \le t \le I$. For a ball $B_i \in \mathcal{H}_1$, let $F(L_t, B_i)$ be the amount of flow rerouted from B_i to L_t. Let $F_t = \sum_{B_i \in \mathcal{H}_1} F(L_t, B_i)$. The next lemma states that when L_t is added to $\mathcal{O}$ sufficient amount of flow is rerouted from the balls in $\mathcal{H}_1$ to L_t irrespective of the value of k_t.

▶ **Lemma 6.** *For $1 \le i \le I$, $F_t \ge k_t/60$ for $\alpha \le 1/60$.*

Proof. To compute the flow rerouted from balls in $\mathcal{H}_1$ to B_t we refer to the three cases mentioned in Cluster Formation. In the first case, for $\lfloor (1-10\alpha)k_t \rfloor$ points, the flow is rerouted from $\mathcal{H}_1$ to B_t. Note that by the invariant we maintain, for each such point p_ℓ, p_ℓ receives at least $1-4\alpha$ amount of flow from the balls in $\mathcal{H}_1$. It follows that, at least $1-4\alpha$ amount of flow is rerouted for p_ℓ and $F_t \ge (1-4\alpha)\lfloor (1-10\alpha)k_t \rfloor \ge (14/15)\lfloor (5/6)k_t \rfloor \ge (14/15)(1/14)k_t = k_t/15 \ge k_t/60$. The second inequality follows as $\alpha \le 1/60$ and the third inequality follows as $k_t > 2$.

In the second case, using the same argument as above, the amount of flow rerouted from $\mathcal{H}_1$ to B_t is at least $1-14\alpha$. As $k_t \le 2$, F_t is at least $(1-14\alpha)k_t/2 \ge (23/60)k_t \ge k_t/60$. The first inequality is true for $\alpha \le 1/60$.

In the third case, again using the same argument as above, the amount of flow rerouted from $\mathcal{H}_1$ to B_t is at least $\min\{(1-10\alpha)k_t - 4\alpha, 1-4\alpha\}$. As $k_t < 1$, $1-4\alpha \ge (1-4\alpha)k_t$. Thus, $F_t \ge (1-10\alpha)k_t - 4\alpha$. As $U_t \ge 1$, $k_t = U_t' \ge 1/10$, and hence $F_t \ge (1-10\alpha)/10 - 4\alpha = 1/10 - 5\alpha \ge 1/60$. The last inequality follows from the fact that $\alpha \le 1/60$. ◀

Define the y-credit of a ball $B_i \in \mathcal{H}_1$ as $Y(L_t, B_i) = F(L_t, B_i)/k_t$. At any moment during the Cluster Formation stage, define the y-accumulation of B_i as $\tilde{y}(B_i) = \sum_{L_t \in \mathcal{O}} Y(L_t, B_i) - \sum_{B_i \in \mathcal{L}_1 \cap \mathrm{cluster}(B_i)} \overline{y}_i$. The y-credit $Y(L_t, B_i)$ of B_i can be seen as a normalized load it transfers to L_t. The y-accumulation $\tilde{y}(B_i)$ is basically the difference between the total y-credit received by B_i and the sum of normalized flows of the balls absorbed by B_i. The next lemma gives a lower bound on the available capacities of the balls in $\mathcal{H}_1$, which is similar to Lemma 3.3 of [4].

▶ **Lemma 7.** *Consider a ball $B_i \in \mathcal{H}_1$ and any integer $1 \leq t \leq I$. Suppose the balls $L_1, \ldots, L_t$ have been added to $\mathcal{O}$ so far. Then, $AC(B_i) \geq \tilde{y}(B_i)k_t$.*

Proof. For any ball $B_i \in \mathcal{H}_1$, we prove the claim using induction on iteration number. In the base case, just after addition of L_1, $AC(B_i) \geq F(L_1, B_i) = Y(L_1, B_i)k_1 = \tilde{y}(B_i)k_1$. Now, suppose the claim is true for any $t-1$. We show that the claim is true for t as well.

Consider the iteration t. Note that $AC(B_i) \geq \tilde{y}(B_i)k_{t-1}$. Suppose a subset of balls have joined cluster of B_i. Let B_p be the first ball joined, which serves k points. To distinguish between the old and new value of $\tilde{y}(B_i)$, we refer to the new value by $\tilde{y}(B_i)'$. After B_p's joining to cluster of B_i, $\tilde{y}(B_i)' = \tilde{y}(B_i) - \overline{y}_p$. Now, the total flow out of B_p is at most $\min\{\overline{y}_p k, \overline{y}_p U_p'\} = \overline{y}_p \min\{k, U_p'\} \leq \overline{y}_p k_{t-1}$. Thus, $AC(B_i) \geq \tilde{y}(B_i)k_{t-1} - \overline{y}_p k_{t-1} = \tilde{y}(B_i)' k_{t-1}$. Using the same argument it can be shown that after each subsequent addition of a ball to cluster of B_i the claim is true.

In the next step, L_t is added to $\mathcal{O}$. Let $\tilde{y}(B_i)$ be the y-accumulation before this. After this addition, the new y-accumulation $\tilde{y}(B_i)' = \tilde{y}(B_i) + Y(L_t, B_i)$. If $\tilde{y}(B_i) \leq 0$, the new available capacity $A_i' \geq Y(L_t, B_i)k_t \geq \tilde{y}(B_i)'k_t$. Otherwise, $\tilde{y}(B_i) > 0$, the new available capacity by the induction hypothesis is, $A_i' = AC(B_i) + Y(L_t, B_i)k_t \geq \tilde{y}(B_i)k_{t-1} + Y(L_t, B_i)k_t \geq (\tilde{y}(B_i) + Y(L_t, B_i))k_t = \tilde{y}(B_i)'k_t$. ◀

The next lemma shows that for any ball $B_i \in \mathcal{H}_1$, y-accumulation is bounded, which is similar to Lemma 3.4 of [4].

▶ **Lemma 8.** *At any point, for any ball $B_i \in \mathcal{H}_1$, $\tilde{y}(B_i) < 1 + 10\alpha$.*

Intuitively, if the y-accumulation of B_i exceeds the bound, it must be due to selection of a ball L_t in $\mathcal{L}_1$. However, one can show that B_i had enough available capacity to absorb the flow from L_t. Hence, the bound follows.

The following lemma gives an upper bound on the number of balls of $\mathcal{L}_1$ that are fully opened.

▶ **Lemma 9.** *At the end of the Cluster Formation stage, $|\mathcal{O}| \leq 60((1+10\alpha)|\mathcal{H}_1| + \sum_{B_i \in \mathcal{L}_1} \overline{y}_i)$.*

Proof.

$$\begin{aligned}
\sum_{B_i \in \mathcal{H}_1} \tilde{y}(B_i) &= \sum_{B_i \in \mathcal{H}_1} \sum_{L_t \in \mathcal{O}} Y(L_t, B_i) - \sum_{B_i \in \mathcal{H}_1} \sum_{B_i \in \mathcal{L}_1 \cap \text{cluster}(B_i)} \overline{y}_i \\
&\geq \sum_{B_i \in \mathcal{H}_1} \sum_{L_t \in \mathcal{O}} F(L_t, B_i)/k_t - \sum_{B_i \in \mathcal{L}_1} \overline{y}_i \\
&= \sum_{t=1}^{I} F_t/k_t - \sum_{B_i \in \mathcal{L}_1} \overline{y}_i \\
&\geq |\mathcal{O}|/60 - \sum_{B_i \in \mathcal{L}_1} \overline{y}_i \qquad (F_t \geq k_t/60 \text{ by Lemma 6})
\end{aligned}$$

Also, by Lemma 8, $\sum_{B_i \in \mathcal{H}_1} \tilde{y}(B_i) \leq (1 + 10\alpha)|\mathcal{H}_1|$. It follows that, $|\mathcal{O}| \leq 60((1 + 10\alpha)|\mathcal{H}_1| + \sum_{B_i \in \mathcal{L}_1} \overline{y}_i)$. ◀

We obtain the following bound on the cost of the rounded solution.

▶ **Lemma 10.** *When the algorithm terminates the total cost of the solution is at most $10|\mathcal{H}_1| + |\mathcal{O}| \leq (70 + 600\alpha)\text{cost}(\sigma^*)/\alpha$.*

Proof. We note that from a heavy balls' cluster at most 10 balls are selected and all the balls in $\mathcal{O}$ are selected. Now, by Lemma 9,

$$\begin{aligned}10|\mathcal{H}_1| + |\mathcal{O}| &\leq 10|\mathcal{H}_1| + 60((1+10\alpha)|\mathcal{H}_1| + \sum_{B_i \in \mathcal{L}_1} \overline{y}_i)\\ &\leq (70+600\alpha)(|\mathcal{H}_1| + \sum_{B_i \in \mathcal{L}_1} \overline{y}_i)\\ &\leq (70+600\alpha)\mathrm{cost}(\sigma^*)/\alpha\end{aligned}$$ ◀

The following lemma shows that 5 factor expansion is sufficient to serve the points assigned to each cluster.

▶ **Lemma 11.** *Using factor 5 expansion of the balls the flow of any cluster can be assigned to the chosen balls without violating the capacities.*

Proof. It is clear from the algorithm that the coverage constraints are satisfied by expanding the balls by at most a factor of 5. Here we consider the capacity constraints. Note that in the first case the capacities of the selected light balls are trivially satisfied. Also, the remaining flow assigned to B_h must have an amount at most U_h due to the way balls are added to a cluster. Thus, its capacity constraint is satisfied. In the other case, let the total amount of flow rerouted from the selected 10 light balls to B_h in Cluster Formation step be f. Also, let B_ℓ be the smallest radius ball among these 10 balls. Thus, the available capacity of all these balls is at least $10U'_\ell - f$. Note that $U_h \leq U_\ell$, as B_ℓ is larger than B_h. Now, as each light balls' capacity is reduced to a factor 10 of the original capacity and the capacity of B_h remains unchanged, $U_h \leq 10U'_\ell$. Hence, the available capacity of all these 10 balls is at least $U_h - f$. As the remaining flow is at most $U_h - f$, it follows that the capacity constraints of these balls are satisfied. ◀

We summarize our findings in the following lemma.

▶ **Lemma 12.** *The solution $(\overline{x}, \overline{y})$ satisfies all the Constraints of AUX-LP1 except Constraint 11. Moreover,*
1. $\overline{y}_i = 1$ *for all* $B_i \in \mathcal{H}_1 \cup \mathcal{O}$ *and* $\overline{y}_i = 0$ *for all other balls.*
2. *For any* $p_j \in P_1$, $\sum_{B_i \in \mathcal{H}_1 \cup \mathcal{O}} \overline{x}_{ij} = 1$.
3. *For any point* $p_j \in P_1$, *if* $\overline{x}_{ij} > 0$, $d(c_i, p_j) \leq 5 \cdot r_i$.
4. $\mathrm{cost}((\overline{x}, \overline{y})) \leq (70+600\alpha)\mathrm{cost}(\sigma^*)/\alpha$.

4.2 Rounding the Second Auxilliary LP

Note that we are given a feasible LP solution $\hat{\sigma} = (\hat{x}, \hat{y})$ to AUX-LP2 that has the following properties.
1. For any $B_i \in \mathcal{L}_2$, $\hat{y}_i \leq 10\alpha$.
2. For any $p_j \in P_2$, $\sum_{B_i \in \mathcal{L}_2} \hat{x}_{ij} \geq 4\alpha$.
3. For any $p_j \in P_2$ and $B_i \in \mathcal{L}_2$, $\hat{x}_{ij} \leq \alpha$.
4. $\mathrm{cost}(\hat{\sigma}) \leq 10 \cdot \mathrm{cost}(\sigma^*)$.

First, we create a new solution to AUX-LP2 from $\hat{\sigma}$ which has cost at most two times that of $\hat{\sigma}$. We denote the new solution as well by $\hat{\sigma}$. Thus, for distinction, we denote the old values by $\hat{y}'_i$ and $\hat{x}'_{ij}$. For each y variable, its new value is twice the old value. Thus, $\hat{y}_i = 2\hat{y}'_i \leq 20\alpha < 1$. The last inequality follows for $\alpha \leq 1/60$. And, for each x variable, its new value is twice the old value. Thus, $\hat{x}_{ij} = 2\hat{x}'_{ij} \leq 2\alpha$. Note that, now, some points might receive flow of more than 1. We adjust the $\hat{x}$ values of these points so that each such point receives 1 amount of flow. We obtain the following lemma.

▶ **Lemma 13.** *There is a feasible LP solution $\hat{\sigma} = (\hat{x}, \hat{y})$ to AUX-LP2 that has the following properties.*

1. *For any $B_i \in \mathcal{L}_2$, $\hat{y}_i \le 20\alpha$.*
2. *For any $p_j \in P_2$, $\sum_{B_i \in \mathcal{L}_2} \hat{x}_{ij} \ge 8\alpha$.*
3. *For any $p_j \in P_2$ and $B_i \in \mathcal{L}_2$, $\hat{x}_{ij} \le 2\alpha$.*
4. $\mathrm{cost}(\hat{\sigma}) \le 20 \cdot \mathrm{cost}(\sigma^*)$.

Proof. First note that $\mathrm{cost}(\hat{\sigma}) \le 20 \cdot \mathrm{cost}(\sigma^*)$, as the values of the y variables are doubled. Next, we show that $\hat{\sigma}$ is feasible.

As the y variables are doubled and $\hat{x}_{ij} \le 2\hat{x}'_{ij}$, $\hat{x}_{ij} \le \hat{y}_i$. Thus, Constraint 14 is satisfied.

For $B_i \in \mathcal{L}_2$, $\sum_{p_j \in P_2} \hat{x}_{ij} \le \sum_{p_j \in P_2} 2\hat{x}'_{ij} = 2\sum_{p_j \in P_2} \hat{x}'_{ij} \le 2\hat{y}'_i \cdot U'_i = \hat{y}_i \cdot U'_i$. Thus, Constraint 15 is satisfied.

As we do not decrease the x variables, unless a point gets more than 1 amount of flow, Constraint 16 is also satisfied. Also, it is trivial to verify that Constraints 17-19 are also satisfied.

Properties 1, 3, and 4 follows immediately. Also, Property 2 follows from the fact that previously each point received a flow of at least 4α from the balls in $\mathcal{L}_2$. Hence, the lemma follows. ◀

We start with the fractional solution $\hat{\sigma} = (\hat{x}, \hat{y})$ and round it so that $\hat{y}$ becomes integral. Throughout our algorithm we modify $\hat{\sigma}$ over several steps to finally obtain the desired solution. Thus whenever we refer to $\hat{\sigma}$ we refer to its current value. For any $p_j \in P_2$, let $\delta_j = \sum_{B_i \in \mathcal{L}_2} \hat{x}_{ij}$. Note that $\delta_j \ge 8\alpha$. Let S and $\mathcal{O}'$ be two disjoint sets of balls which are initialized to $\mathcal{L}_2$ and $\emptyset$, respectively. Throughout we also maintain that $\sum_{B_i \in S \cup \mathcal{O}'} \hat{x}_{ij} = \delta_j$. Note that this is true in the beginning. Our algorithm is as follows.

While there is a point $p_j \in P_2$ such that $\sum_{B_i \in S} \hat{x}_{ij} > \alpha$, we do the following.

Let S_j be the set of balls in S that give flow to p_j, i.e., S_j=$\{B_i \in S : \hat{x}_{ij} > 0\}$. Note that as $\sum_{B_i \in S_j} \hat{x}_{ij} = \sum_{B_i \in S} \hat{x}_{ij} > \alpha$, $\sum_{B_i \in S_j} \hat{y}_i \ge \sum_{B_i \in S_j} \hat{x}_{ij} > \alpha$. Find $T \subseteq S_j$ such that $\alpha \le \sum_{B_i \in T} \hat{y}_i \le 21\alpha$. Such a subset can always be found using a linear scan of S_j, as $\sum_{B_i \in S_j} \hat{y}_i > \alpha$ and $\hat{y}_i \le 20\alpha$ for all $B_i \in S_j$. Let B_t be the largest ball in T. Set $\hat{y}_t = 1$ and $\hat{y}_i = 0$ for each $B_i \in T$. Add B_t to $\mathcal{O}'$. Remove all balls in T from S. Reroute the flow from all balls in $T \setminus \{B_t\}$ to B_t.

▶ **Lemma 14.** *During the course of the above algorithm, the solution $\hat{\sigma}$ has cost at most $20 \cdot \mathrm{cost}(\sigma^*)/\alpha$ and satisfies all the constraints of AUX-LP2 except Constraint 17. Moreover, for a point $p_j \in P_2$, if $\hat{x}_{ij} > 0$, $d(c_i, p_j) \le 3 \cdot r_i$.*

Proof. First, we prove the feasibility of $\hat{\sigma}$ using induction on the iteration number. In the beginning, the claim holds. Now, consider a particular iteration. Note that the balls for which the $\hat{y}$ values are changed are in T and the points for which the $\hat{x}$ values are changed are the set of points P' that receive flow from a ball in T. It is sufficient to show that the constraints concerning these balls and points hold. Constraint 14 is satisfied as for each such point p_j, and the ball B_t, $\hat{x}_{tj} \le \delta_j \le 1 = \hat{y}_t$ and for a ball $B_i \in T \setminus \{B_t\}$, $\hat{x}_{ij} = 0$. Now, we argue that the capacity constraint of the ball B_t is satisfied. Note that in the beginning of the iteration, the total flow out of balls in T to all points is at most

$$\sum_{B_i \in T} \hat{y}_i \cdot U'_i \le U'_t \sum_{B_i \in T} \hat{y}_i \le U'_t \cdot 21\alpha < U'_t.$$

The first inequality follows from the fact that B_t is the largest ball in T and all the capacities of the balls in $\mathcal{L}_1$ are scaled by the same factor. The last inequality follows, as $\alpha \le 1/60$.

Now, as this total flow is served by B_t the claim holds. Constraint 16 is also satisfied for all the points in P', as the flow is only rerouted from a ball to B_t. The other constraints except 17 are trivial to verify.

Note that whenever we set $\hat{y}_t = 1$, we also set $\hat{y}_i = 0$ for each $B_i \in T \setminus \{B_t\}$. Thus for each ball B_t we can charge all the balls in T. As $\sum_{B_i \in T} \hat{y}_i \geq \alpha$, the cost blow up is at most a factor of $1/\alpha$. Thus, the cost is at most $20 \cdot \text{cost}(\sigma^*)/\alpha$.

Whenever we reassign flow from balls in $T \setminus \{B_t\}$ to B_t, for a point $p_j \in P_2$, it holds that if $\hat{x}_{tj} > 0$, $d(c_t, p_j) \leq 3 \cdot r_t$. This is true, as B_t is the largest ball in T. As we remove B_t from S, no flow is ever rerouted again from or to B_t. Hence, the claim continues to hold for all points. ◀

Now, note that when the while loop of the above algorithm terminates, it holds that for any $p_j \in P_2$, $\sum_{B_i \in S} \hat{x}_{ij} \leq \alpha$. Thus, $\sum_{B_i \in \mathcal{O}'} \hat{x}_{ij} \geq \delta_j - \alpha \geq 7\alpha$. Using this fact, we compute a solution (x', y') to AUX-LP2 (that violates Constraint 17 and Constraint 15). For any ball B_i in $\mathcal{O}'$, set $y'_i = 1$. For any $p_j \in P_2$ and B_i in $\mathcal{O}'$, set $x'_{ij} = \min\{(1/(7\alpha)) \cdot \hat{x}_{ij}, 1\}$. All the other x' and y' values are set to zero. Note that, now, each point receives a flow of at least 1. We adjust the x' values so that each point receives exactly 1 amount of flow. We obtain the following lemma.

▶ **Lemma 15.** *The solution (x', y') satisfies all the constraints of AUX-LP2 except Constraint 17 and Constraint 15. Moreover,*

1. $y'_i = 1$ *for all* $B_i \in \mathcal{O}'$ *and* $y'_i = 0$ *for all* $B_i \notin \mathcal{O}'$.
2. *For any* $p_j \in P_2$, $\sum_{B_i \in \mathcal{O}'} x'_{ij} = 1$.
3. *For any* $B_i \in \mathcal{O}'$, $\sum_{p_j \in P_2} x'_{ij} \leq (1/(7\alpha)) \cdot U'_i$.
4. *For any point* $p_j \in P_2$, *if* $x'_{ij} > 0$, $d(c_i, p_j) \leq 3 \cdot r_i$.
5. $\text{cost}((x', y')) \leq 20 \cdot \text{cost}(\sigma^*)/\alpha$.

4.3 Combining the Two LP solutions

Next, we compose the two rounded solutions obtained in Lemma 12 and 15 to construct a solution for the original instance. In the new solution $(\tilde{x}, \tilde{y})$ we fully open the balls in $\mathcal{H}_1 \cup \mathcal{O} \cup \mathcal{O}'$. Also we keep all the x values unchanged. Note that a ball B_i of $\mathcal{L}_1 (= \mathcal{L}_2)$ can be opened in both solutions. However, as we had changed its capacity before, the total capacity that it can use is at most $U'_i + (1/(7\alpha)) \cdot U'_i \leq (1 + 1/(7\alpha))U_i/10 < U_i$. The last inequality follows by setting $\alpha = 1/60$. The total cost of the new solution is at most $(90 + 600\alpha)\text{cost}(\sigma^*)/\alpha \leq 6000 \cdot \text{cost}(\sigma^*)$. Hence, we obtain the following lemma.

▶ **Lemma 16.** *The solution $(\tilde{x}, \tilde{y})$ satisfies all the Constraints of MMCC-LP except Constraint 11. Moreover,*

1. *For any point* $p_j \in P_1$, *if* $\overline{x}_{ij} > 0$, $d(c_i, p_j) \leq 5 \cdot r_i$.
2. $\text{cost}((\tilde{x}, \tilde{y})) \leq 6000 \cdot \text{cost}(\sigma^*)$.

We note that by selecting different values of the parameters throughout the algorithm one can improve the constant in the approximation factor. However, as our main goal is to show any $O(1)$-approximation we did not pursue this.

▶ **Theorem 17.** *There is an $O(1)$-approximation for MMCC by expanding the balls by a factor of at most* 5.

5 Uniform Capacitated Case

The algorithm in the uniform case is same except the Selection of Balls step. The next lemma shows that the Selection of Balls can be performed with only 4.24 factor expansion of the balls.

▶ **Lemma 18.** *Using factor* 4.24 *expansion of the balls the flow of any cluster can be assigned to the chosen balls without violating the capacities.*

Proof. Consider any cluster of a heavy ball $B_h \in \mathcal{H}_1$. Let $c = (1+\sqrt{5})/2$. If B_h is one of the top 10 largest balls in the cluster, then select all the balls larger than B_h and also B_h. Only B_h is expanded by a factor of 3. The flow rerouted from any selected ball of $\mathcal{L}_1$ to B_h is assigned to the selected ball. Note that for the remaining balls of $\mathcal{L}_1$ which are in the same cluster and not chosen, are smaller than B_h and thus can be covered by a factor 3 expansion of B_h. The remaining flow is assigned to B_h. Note that in this case the capacities of the selected light balls are trivially satisfied. Also, the remaining flow assigned to B_h must have an amount at most U_h. Thus, the capacity constraint of B_h is satisfied.

Now, suppose B_h is not one of the top 10 largest balls. Let B_ℓ be the 10^{th} largest ball of this cluster. Also, let r_h and r_ℓ be the radius of B_h and B_ℓ, respectively. Now, there can be two cases (i) $r_h \geq r_\ell/c$ or (ii) $r_h < r_\ell/c$. In the first case, we select the top 9 largest balls all of which are in $\mathcal{L}_1$ and also B_h. The flow rerouted from any selected ball (except B_h) to B_h is assigned to the selected ball. Now consider the remaining flow assigned to the cluster. Also consider a point p_j which receives a part of this flow and not in any of the balls selected from $\mathcal{L}_1$. Then, by triangle inequality, the distance between p_j and the center c_h of B_h is at most $r_h + 2r_\ell \leq r_h + 2cr_h \leq 4.24r_h$. We expand B_h by the factor 4.24 and assign the remaining flow to B_h. Selected balls which are in $\mathcal{L}_1$ are not expanded. The capacity constraints are also satisfied due to the same reason mentioned above.

In the second case, the top 10 largest balls are selected all of which are in $\mathcal{L}_1$. The flow rerouted from any selected ball to B_h is assigned to the selected ball. Now consider the remaining flow assigned to the cluster. Also consider a point p_j which receives a part of this flow and not in any of the selected balls. Let $B_t = B(c_t, r_t)$ be a selected ball. Then, by triangle inequality, the distance between p_j and c_t is at most $r_t+2r_h+2r_\ell \leq r_t+2r_\ell/c+2r_\ell \leq (3+2/c)r_t \leq 4.24r_t$. The second last inequality follows, as r_ℓ is the smallest of the selected balls. We expand each selected ball by the factor 4.24. The remaining flow is assigned arbitrarily to selected balls respecting their capacity. Let the total amount of flow rerouted from the selected 10 light balls to B_h in Cluster Formation step be f. The total available capacity of all these balls is at least $10U'_\ell - f$, as B_ℓ is the smallest radius ball among these 10 balls. Now, as the capacity of each ball of $\mathcal{L}_1$ is reduced to a factor 10 of the original capacity and the capacity of B_h remains unchanged, $U_h \leq 10U'_\ell$. Hence, the available capacity of all these 10 balls is at least $U_h - f$. As the remaining flow is at most $U_h - f$, it follows that the capacity constraints of these balls are satisfied. ◀

▶ **Theorem 19.** *There is an* $O(1)$*-approximation for MCC by expanding the balls by a factor of at most* 4.24*.*

6 Conclusion

In this paper, we improve the expansion factor of the balls for MCC and MMCC to 4.24 and 5, respectively, in the context of obtaining constant approximation. Our approximation factor is a large constant. But, it is possible to improve this factor by setting different values of parameters in the algorithm. Note that the lower bound on the expansion factor is still 3. So, one obvious problem is to reduce the gap further. Another interesting problem is to design a true constant approximation for the Euclidean version of MCC, which does not expand the balls. We note that this problem is open even in the plane.

Note that if the capacities are not monotonic, no $(O(1), O(1))$-approximation is known. On the other hand, the lower bound on the expansion factor even in this case is $3 - \epsilon$, similar to the uniform capacity case. So, a very natural and interesting direction of research is to study this most general version of the problem.

References

1. Ankit Aggarwal, Anand Louis, Manisha Bansal, Naveen Garg, Neelima Gupta, Shubham Gupta, and Surabhi Jain. A 3-approximation algorithm for the facility location problem with uniform capacities. *Math. Program.*, 141(1-2):527–547, 2013.
2. Hyung-Chan An, Aditya Bhaskara, Chandra Chekuri, Shalmoli Gupta, Vivek Madan, and Ola Svensson. Centrality of trees for capacitated k-center. *Math. Program.*, 154(1-2):29–53, 2015. doi:10.1007/s10107-014-0857-y.
3. Hyung-Chan An, Mohit Singh, and Ola Svensson. Lp-based algorithms for capacitated facility location. *SIAM J. Comput.*, 46(1):272–306, 2017.
4. Sayan Bandyapadhyay, Santanu Bhowmick, Tanmay Inamdar, and Kasturi Varadarajan. Capacitated covering problems in geometric spaces. *Discrete & Computational Geometry*, pages 1–31, 2019.
5. Manisha Bansal, Naveen Garg, and Neelima Gupta. A 5-approximation for capacitated facility location. In Leah Epstein and Paolo Ferragina, editors, *Algorithms - ESA 2012 - 20th Annual European Symposium, Ljubljana, Slovenia, September 10-12, 2012. Proceedings*, volume 7501 of *Lecture Notes in Computer Science*, pages 133–144. Springer, 2012.
6. Judit Bar-Ilan, Guy Kortsarz, and David Peleg. How to allocate network centers. *J. Algorithms*, 15(3):385–415, 1993. doi:10.1006/jagm.1993.1047.
7. Hervé Brönnimann and Michael T. Goodrich. Almost optimal set covers in finite vc-dimension. *Discrete & Computational Geometry*, 14(4):463–479, 1995.
8. Jaroslaw Byrka, Krzysztof Fleszar, Bartosz Rybicki, and Joachim Spoerhase. Bi-factor approximation algorithms for hard capacitated k-median problems. In Piotr Indyk, editor, *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 722–736. SIAM, 2015.
9. Jaroslaw Byrka, Bartosz Rybicki, and Sumedha Uniyal. An approximation algorithm for uniform capacitated k-median problem with 1+\epsilon capacity violation. In Quentin Louveaux and Martin Skutella, editors, *Integer Programming and Combinatorial Optimization - 18th International Conference, IPCO 2016, Liège, Belgium, June 1-3, 2016, Proceedings*, volume 9682 of *Lecture Notes in Computer Science*, pages 262–274. Springer, 2016.
10. Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the k-median problem. *J. Comput. Syst. Sci.*, 65(1):129–149, 2002.
11. Fabián A. Chudak and David P. Williamson. Improved approximation algorithms for capacitated facility location problems. *Math. Program.*, 102(2):207–222, 2005.
12. Julia Chuzhoy and Joseph Naor. Covering problems with hard capacities. *SIAM J. Comput.*, 36(2):498–515, 2006.
13. Julia Chuzhoy and Yuval Rabani. Approximating k-median with non-uniform capacities. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005*, pages 952–958. SIAM, 2005.
14. Marek Cygan, MohammadTaghi Hajiaghayi, and Samir Khuller. LP rounding for k-centers with non-uniform hard capacities. In *FOCS*, pages 273–282, 2012.

15 H. Gökalp Demirci and Shi Li. Constant approximation for capacitated k-median with (1+epsilon)-capacity violation. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 73:1–73:14, 2016.

16 Uriel Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, 1998.

17 Rajiv Gandhi, Eran Halperin, Samir Khuller, Guy Kortsarz, and Srinivasan Aravind. An improved approximation algorithm for vertex cover with hard capacities. *J. Comput. Syst. Sci.*, 72(1):16–33, 2006.

18 Taha Ghasemi and Mohammadreza Razzazi. A PTAS for the cardinality constrained covering with unit balls. *Theor. Comput. Sci.*, 527:50–60, 2014.

19 Sariel Har-Peled and Mira Lee. Weighted geometric set cover problems revisited. *JoCG*, 3(1):65–85, 2012.

20 Mong-Jen Kao. Iterative partial rounding for vertex cover with hard capacities. In *SODA*, pages 2638–2653, 2017.

21 Samir Khuller and Yoram J. Sussmann. The capacitated K-center problem. *SIAM J. Discrete Math.*, 13(3):403–418, 2000.

22 Madhukar R. Korupolu, C. Greg Plaxton, and Rajmohan Rajaraman. Analysis of a local search heuristic for facility location problems. *J. Algorithms*, 37(1):146–188, 2000.

23 Shi Li. On uniform capacitated k-median beyond the natural LP relaxation. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 696–707, 2015.

24 Shi Li. On uniform capacitated k-median beyond the natural LP relaxation. *ACM Trans. Algorithms*, 13(2):22:1–22:18, 2017.

25 Robert Lupton, F. Miller Maley, and Neal E. Young. Data collection for the sloan digital sky survey - A network-flow heuristic. *J. Algorithms*, 27(2):339–356, 1998. `doi:10.1006/jagm.1997.0922`.

26 Nabil H. Mustafa and Saurabh Ray. Improved results on geometric hitting set problems. *Discrete & Computational Geometry*, 44(4):883–895, 2010.

27 Martin Pál, Éva Tardos, and Tom Wexler. Facility location with nonuniform hard capacities. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 329–338. IEEE Computer Society, 2001.

28 Laurence A. Wolsey. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, 2(4):385–393, 1982.

29 Sam Chiu-wai Wong. Tight algorithms for vertex cover with hard capacities on multigraphs and hypergraphs. In *SODA*, pages 2626–2637, 2017.