Improved Approximation Algorithm for Set Multicover with Non-Piercing Regions

Rajiv Raman
IIIT Delhi, India
rajiv@iiitd.ac.in

Saurabh Ray
NYU Abu Dhabi, United Arab Emirates
saurabh.ray@nyu.edu

Abstract

In the Set Multicover problem, we are given a set system $(X, \mathcal{S})$, where X is a finite ground set, and $\mathcal{S}$ is a collection of subsets of X. Each element $x \in X$ has a non-negative demand $d(x)$. The goal is to pick a smallest cardinality sub-collection $\mathcal{S}'$ of $\mathcal{S}$ such that each point is covered by at least $d(x)$ sets from $\mathcal{S}'$. In this paper, we study the set multicover problem for set systems defined by points and non-piercing regions in the plane, which includes disks, pseudodisks, k-admissible regions, squares, unit height rectangles, homothets of convex sets, upward paths on a tree, etc.

We give a polynomial time $(2+\epsilon)$-approximation algorithm for the set multicover problem $(P, \mathcal{R})$, where P is a set of points with demands, and $\mathcal{R}$ is a set of non-piercing regions, as well as for the set multicover problem $(\mathcal{D}, P)$, where $\mathcal{D}$ is a set of pseudodisks with demands, and P is a set of points in the plane, which is the hitting set problem with demands.

2012 ACM Subject Classification Theory of computation → Packing and covering problems; Theory of computation → Computational geometry

Keywords and phrases Approximation algorithms, geometry, Covering

Digital Object Identifier 10.4230/LIPIcs.ESA.2020.78

Acknowledgements The authors are grateful to Nabil H. Mustafa for ideas and comments.

1 Introduction

The Set Cover problem and its variants are central problems in Computer Science. For general set systems, tight results are known - there is an $O(\log n)$ approximation algorithm [32] and this is tight under standard complexity assumptions [18]. Over the last decade, significant progress has been on these problems for geometric set systems in the plane and in low dimensions. There are broadly two approaches that have been successful in the geometric setting, *viz.*, LP-rounding based algorithms and local-search. The LP-rounding approach relies on the existence of small ϵ-nets, which exist whenever the *VC-dimension* of the set system is bounded. Set systems with VC-dimension at most d have ϵ-nets of size $O(d/\epsilon \log 1/\epsilon)$ [22], and this leads to an $O(\log |\textsc{Opt}|)$-approximation algorithm [5, 17], that holds even in the *weighted setting*[1]. Smaller ϵ-nets are known when the *union complexity* of the set system is *small* [13], leading to algorithms with better approximation factors, though not in the weighted setting. Varadarajan [31] showed via the *quasi-uniform* sampling technique how these results can be made to work in the weighted setting. His technique was optimized by Chan, et al. [7] who also introduced the notion of *shallow cell complexity* generalizing the notion of union complexity to abstract set systems. Chekuri, et al. [11] extended the LP-based techniques for set cover to set multicover obtaining an $O(\log |\textsc{Opt}|)$

[1] In the weighted setting, the sets have non-negative weights, and the goal is to find a minimum weight feasible sub-collection, as opposed to the minimum cardinality.

28th Annual European Symposium on Algorithms (ESA 2020).
Editors: Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders; Article No. 78; pp. 78:1–78:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

approximation when the VC dimension is bounded, and better bounds in the case where the union complexity is bounded. In particular, they obtain $O(1)$-approximation algorithms when union complexity is linear. However, their results only hold in the unweighted setting. Bansal and Pruhs [3] extended the approach based on shallow cell complexity and quasi-uniform sampling [31, 7] to work for the weighted set multicover problem. The main weakness of these LP-rounding techniques is that the approximation factor obtained is at least as large as the integrality gap, which is often large. Furthermore, the constants gained in the approximation factor during the rounding process is often large. For instance, [11] uses *shallow cuttings*, which involves large constants.

A second approach that has been effective for fundamental geometric packing and covering problems, albeit in the unweighted setting [8, 28, 21, 25, 4, 30] is Local Search. Besides packing and covering problems, Local Search has also been remarkably successful for several clustering problems (See [20, 19, 14, 15], and references therein). A drawback of the Local Search approach is that the running time of the algorithms are prohibitive. In particular, Mustafa and Jartoux [23] showed that to obtain a $(1+\epsilon)$-approximation for the Set Cover problem with disks, the local search algorithm takes $n^{\Omega(1/\epsilon^2)}$ time.

The analysis of local search for most of the geometric packing and covering problems relies on showing the existence of a graph with desired characteristics which is problem specific, and this is usually the challenging part of the analysis. In [30], Raman and Ray gave a unified method to obtain such graphs for several packing and covering problems. While the techniques in [30] can be extended to *packing problems* with bounded capacities as was shown in [4], it was not clear how to extend them to the set multicover problem - even with bounded demands.

Obtaining approximation algorithms that run fast, while simultaneously guaranteeing small approximation factors is a challenging research direction. Recently, Chekuri et al. [12], and Chan and He [10], building on the work of Agarwal and Pan [1] have improved the running times of the LP based algorithms for both packing and covering problems via the multiplicative weights update framework. In this work, we improve the approximation factor, but we do not improve the running times of the algorithms. We give a polynomial time $(2+\epsilon)$-approximation algorithm for the Set Multicover problem for non-piercing regions in the unweighted setting. We obtain the same approximation factor for the multi-hitting set problem for pseudodisks. Our key observation is that even if the LP relaxation has a large integrality gap, we can round it without losing more than a $(1+\delta)$ factor so that it meets all demands with only a constant deficit which depends on the parameter δ. This yields a problem with low demands for which a PTAS can be obtained via local search. Note that even the second part of our approach - PTAS for multihitting set problems with bounded demands, is non-trivial, and builds on tools developed by Raman and Ray [30].

2 Preliminaries

The set multicover problem is defined by a set system $(X, \mathcal{S})$ and a demand function $d : X \mapsto \mathbb{R}$. The task is to select the smallest cardinality subset $\mathcal{S}'$ of $\mathcal{S}$ such that each $x \in X$ is contained in at least $d(x)$ subsets of $\mathcal{S}'$. We refer to the set X as the *ground set* and the elements of $\mathcal{S}$ as *ranges*. For any subset $\mathcal{S}' \subseteq \mathcal{S}$ and any $x \in X$, we denote by $\mathcal{S}'(x)$ the set of elements in $\mathcal{S}'$ containing x and we refer to $|\mathcal{S}'(x)|$ as the *depth* of x in $\mathcal{S}'$.

We require the notion of *shallow-cell complexity* defined by Chan et al. [7]. A *cell* in the set system $(X, \mathcal{S})$ is a maximal subset $X' \subseteq X$ such that the elements of X' are contained in the same collection of ranges in $\mathcal{S}$. We say that a range $S \in \mathcal{S}$ contains a cell C if S contains the elements in C. The depth of a cell C is the number of ranges in $\mathcal{S}$ containing C. A set system has shallow-cell complexity $f(n,k)$ if for any subset $\mathcal{S}' \subseteq \mathcal{S}$ of size n, the number of cells of depth at most k in $(X, \mathcal{S}')$ is at most $f(n,k)$. We focus on set systems whose shallow cell complexity is linear in n and polynomial in k. We say that a set system is *c-linear*, for some constant c, if its shallow-cell complexity is $O(nk^c)$.

The set systems we study in this paper are defined by a set of points and a set of regions in the plane. A *Jordan region* is a compact, simply connected set in the plane whose boundary is a simple jordan curve. We say that a set $\mathcal{R}$ of Jordan regions is a non-piercing family if for any $\gamma, \gamma' \in \mathcal{R}$, $\gamma \setminus \gamma'$ and $\gamma' \setminus \gamma$ are both path connected sets. A set of Jordan regions $\mathcal{R}$ is said to be a family of *pseudodisks* if the boundaries of every pair of regions either do not intersect or cross (i.e. intersect non-tangentially) at exactly two points. Note that a family of pseudodisks is also a family of non-piercing regions but not vice-versa.

In this paper, we study two set multicover problems defined by a set of points and a set of regions in the plane. The first is the set multicover problem in which the ground set is a finite set P of points in the plane and the ranges are obtained by intersecting P with the regions in a family $\mathcal{R}$ of non-piercing regions. Abusing notation, we denote such set systems by $(P, \mathcal{R})$. In the second set multicover problem we study, the ground set is a family of pseudodisks $\mathcal{D}$ and each range is the subset of $\mathcal{D}$ containing a particular point p in a set of points P. Again, for simplicity, we denote such a set system by $(\mathcal{D}, P)$. This variant of the set cover problem is usually called the *hitting set* problem.

3 Our Results

The results of Chekuri et al. [11] imply an $O(1)$-approximation algorithm for the set multicover problem for set systems with linear union complexity. Such set systems are c-linear for some constant c. Bansal and Pruhs [3] guarantee an $O(1)$-approximation factor for the weighted set multicover problem defined by a c-linear set system.

Since both the methods above are based on LP-rounding, the approximation factor is at least as large as the integrality gap. Even for set cover problems (i.e., all demands are 1) defined by very simple geometric regions in the plane, the integrality gap is not known to be small. For halfspaces in the plane, the integrality gap is 2 [24] which implies that it is at least 2 for disks as well. Even though no larger lower bound is known, the best upper bound currently known on the integrality gap for disks is significantly higher: 13.4 [6]. The integrality gap is probably higher when the demands are allowed to be more than 1. However, we are not aware of any results regarding this. Apart from the integrality gap, there are additional constant factors that are gained in the process of rounding. While the exact constants are not analyzed in [11] or [3], they seem to be large. For instance, since the main tool is used in [3] is the quasi-sampling technique [7, 31], the constant seems to be at least $e^{4.34362} > 76$ (see Claims 2 and 4 of [7]). We do not have accurate estimates for the constants involved in shallow cuttings used in [11] but we believe that they are not significantly smaller (and likely to be much larger since they use ϵ-nets and approximations for which the known constants are quite large). Our main result is a $(2+\epsilon)$-approximation algorithm for the set multicover problem for points and non-piercing regions in the plane, in the unweighted setting.

▶ **Theorem 1.** *The set multicover problem defined by a set system $(P, \mathcal{R})$, where P is a finite set of points and $\mathcal{R}$ is a set of non-piercing regions in the plane with an arbitrary demand function $d : P \to \mathbb{R}$ admits a polynomial time $(2+\epsilon)$-approximation algorithm for any $\epsilon > 0$.*

The result follows from the following results that we prove in Sections 4 and 5. First, we show that the set system $(P, \mathcal{R})$ is 2-linear (Lemma 25) and for any set system that is c-linear for some constant c, we obtain the following result via a simple modification of the technique of Bansal and Pruhs [3].

▶ **Theorem 2.** *Let $(X, \mathcal{S})$ be a c-linear set system for some constant c and let $m = |\mathcal{S}|$. Consider the set multicover problem defined by $(X, \mathcal{S})$ in which each element $x \in X$ has a demand $d(x)$. Let y be any feasible solution to the linear programming relaxation of this problem i.e., y satisfies the constraints: $\forall x \in X$, $\sum_{S \ni x} y_S \geq d(x)$, and let $\delta \in (0, 1)$ be a given parameter. Then, we can obtain a subset of ranges $\mathcal{S}' \subseteq \mathcal{S}$ s.t.*

(i) $\forall x \in X$, $|\mathcal{S}'(x)| \geq d(x) - 3\tau$, *where* $\tau = C\delta^{-4} \log \delta^{-1}$ *with a large enough* C

(ii) $|\mathcal{S}'| \leq (1 + O(\delta)) \sum_{S \in \mathcal{S}} y_S$

This shows that the LP solution can be "rounded" without increasing the objective value much and causing only a constant *deficit* in the demands. Since the *residual* demands are $O(1)$ for any constant δ, we obtain a set multicover problem with bounded demands. We then show that the set multicover problem with non-piercing regions and bounded demands has a PTAS via local search.

▶ **Theorem 3.** *Local Search yields a PTAS for the set multicover problem defined by a set system $(P, \mathcal{R})$ where P is a set of points and $\mathcal{R}$ is a family of non-piercing regions in the plane and where each point $p \in P$ has a demand bounded above by some constant Θ.*

We also obtain the same approximation factor in the dual setting, though here we are currently only able to prove the result when the regions are pseudodisks.

▶ **Theorem 4.** *Let $\mathcal{P}$ be a Set Multicover problem for the set system $(\mathcal{D}, P)$ defined by a set of pseudodisks $\mathcal{D}$ and a set of points P in the plane with demand function $d : \mathcal{D} \to \mathbb{R}$. For any $\epsilon > 0$, there is a polynomial time $(2 + \epsilon)$-approximation algorithm for $\mathcal{P}$.*

The result is obtained by showing that the set system $(\mathcal{D}, P)$ is 2-linear, which implies that Theorem 2 can be used to obtain an instance of the set multicover problem where the demands of the pseudodisks are bounded above by a constant Θ. For these instances, we show that local search yields a PTAS.

▶ **Theorem 5.** *Local Search yields a PTAS for the set multicover problem defined by a set system $(\mathcal{D}, P)$ where P is a set of points and $\mathcal{D}$ is a family of pseudodisks in the plane and where each pseudodisk $D \in \mathcal{D}$ has a demand bounded above by some constant Θ.*

Even though the PTASes obtained in Theorems 3 and 5 are not surprising, the proofs are not trivial. In fact, it would not be surprising if local search yields a PTAS for the set multicover problem with arbitrary demands. However, we are currently unable to prove such a result.

The paper is organized as follows. In Section 4, we prove Theorem 2. In Subsection 5.1, we prove Theorem 3, and we prove Theorem 5 in Subsection 5.2. Theorems 1 and 4 are proved in Section 6.

4 LP Rounding with bounded deficits

In this section we prove Theorem 2. Consider the set multicover problem defined by a c-linear set system $(X, \mathcal{S})$ and a demand function $d : X \mapsto \mathbb{R}$. The natural linear programming relaxation for this problem is the following:

$$\min \sum_{S \in \mathcal{S}} y_S \quad \text{s.t.} \quad \forall S \in \mathcal{S}, \; y_S \in [0,1] \quad \text{and} \quad \forall x \in X, \sum_{S \ni x : S \in \mathcal{S}} y_S \geq d(x) \tag{1}$$

Our goal is to show that given any feasible solution y to the above LP and any $\delta > 0$, we can find a subset $\mathcal{S}' \subseteq \mathcal{S}$ of size at most $(1 + O(\delta)) \sum_{S \in \mathcal{S}} y_S$ and $\forall x \in X, |\mathcal{S}'(x)| \geq d(x) - O(\delta^{-4} \log \delta^{-1})$. We start with a few technical results. The lemma below follows from the techniques in [7, 31] where similar statements are proved.

▶ **Lemma 6.** *Let $(X, \mathcal{S})$ be a c-linear set system and let $m = |\mathcal{S}|$. Then, there exists an ordering $S_1, \cdots, S_m$ of the ranges in $\mathcal{S}$ s.t. in the set system $(X, \mathcal{S}_i)$ where $\mathcal{S}_i = \{S_1, \cdots, S_i\}$, the range S_i contains at most $O(k^{c+3})$ cells of depth at most k, for any k.*

Proof. We assign a weight to each cell in $(X, \mathcal{S})$ depending on its depth. A cell of depth k is assigned a weight of $1/k^{c+3}$. We define the weight of a range to be the total weight of all the cells it contains. Since there are at most $O(mk^c)$ cells of depth k (since $(X, \mathcal{S})$ is c-linear), the total weight of all cells of depth k is at most $O(m/k^3)$. Since each cell of depth k contributes to the weight of k ranges in $\mathcal{S}$, the contribution of the depth k cells to the total weight of all ranges is $O(m/k^2)$. The total weight of all ranges is therefore at most $O(m) \sum_{k=1}^{m} k^{-2} = O(m)$. This implies that there is a range $S \in \mathcal{S}$ whose weight is $O(1)$ which in turn implies that for any k, the number of depth k cells in S is at most $O(k^{c+3})$. We recursively find the ordering of $(X, \mathcal{S} \setminus \{S\})$. The ordering for $(X, \mathcal{S})$ is obtained by appending S to the ordering for $(X, \mathcal{S} \setminus \{S\})$. The lemma follows. ◀

▶ **Definition 7** (Weighted Depth). *Let $(X, \mathcal{S})$ be a set system. Let $w(S) \geq 0$ be a weight associated with each range $S \in \mathcal{S}$. Then, for any element $x \in X$, we denote by $w_{\mathcal{S}}(x)$ the total weight of the ranges in $\mathcal{S}$ containing x. We call this the weighted depth of the element x with respect to the ranges in $\mathcal{S}$ and the weight function w.*

▶ **Lemma 8** (Weighted Sampling Procedure). *Let $(X, \mathcal{S})$ be a c-linear set system for some constant c and let $m = |\mathcal{S}|$. Let $w(S) \in [0.5, 1.0]$ be a weight associated with each $S \in \mathcal{S}$ and let $\delta \in (0,1)$ be a parameter. Then, there is a polynomial time procedure to pick a subset $\mathcal{S}' \subseteq \mathcal{S}$ s.t.*
- **(i)** *$|\mathcal{S}'| \leq (1 + O(\delta))\, W$ where $W = \sum_{S \in \mathcal{S}} w(S)$, and*
- **(ii)** *for any element $x \in X$: $|\mathcal{S}'(x)| \geq w_{\mathcal{S}}(x) - \tau$ where $\tau = C\delta^{-3} \log \delta^{-1}$ for a large enough constant C.*

Proof. The set $\mathcal{S}'$ is unweighted but we can think of each range in $\mathcal{S}'$ as having weight 1. This means that any range $S \in \mathcal{S}$ with weight $\geq 1 - \delta$ can be safely included in $\mathcal{S}'$ since this way we are increasing its weight by at most a factor of $1 + O(\delta)$. We will thus assume without loss of generality that all ranges in $\mathcal{S}$ have weight at most $1 - \delta$. We will show that a procedure similar to the quasi-uniform sampling procedure [31, 7] can be used to pick each range $S \in \mathcal{S}$ into $\mathcal{S}'$ with probability at most $(1 + O(\delta))\, w(S)$ s.t. the second condition in the lemma is satisfied. Then, the expected number of ranges in $\mathcal{S}'$ is at most $(1 + O(\delta))\, W$.

Now, by Markov inequality, the probability that $|\mathcal{S}'|$ exceeds its expectation by a factor of more than $(1 + \delta)$ is at most $1/(1 + \delta)$ which means that the with probability $\Omega(\delta)$, $|\mathcal{S}'|$ is $(1 + O(\delta))W$. We can therefore repeat the process $O(1/\delta)$ times in expectation to obtain the desired collection $\mathcal{S}'$ of sets.

Our sampling procedure has two stages. In the first stage, we pick a sample $\mathcal{T} \subseteq \mathcal{S}$ by picking each range $S \in \mathcal{S}$ independently with probability $w(S)/(1-\delta) \in [0,1]$. In the second stage, we pick another sample $\mathcal{T}' \subseteq \mathcal{S}$ s.t. each range is picked with a probability $O(\delta) = O(w(S) \cdot \delta)$ but whether or not a range is picked depends on the outcome of the first stage. We set $\mathcal{S}' = \mathcal{T} \cup \mathcal{T}'$. The probability that a particular range $S \in \mathcal{S}$ is included in $\mathcal{S}'$ is at most $w(S)/(1-\delta) + O(w(S) \cdot \delta) = w(S)(1 + O(\delta))$ as required.

Let $S_1, \cdots, S_m$ be the ordering of the ranges in $\mathcal{S}$ given by Lemma 6. Let $\mathcal{S}_i = \{S_1, \cdots, S_i\}$ and let $\mathcal{T}_i = \mathcal{T} \cap \mathcal{S}_i$ denote the subset of the ranges in $\mathcal{S}_i$ picked in the first stage. The range S_i is picked in the second stage if it is *forced* by an element $x \in X$ which happens if $k = w_{\mathcal{S}}(x) \geq \tau$, $k_i = w_{\mathcal{S}_i}(x) \geq \delta k$ and $s_i = |\mathcal{T}_i(x)| < k_i$. In words, S_i is forced by x if i) x has high weighted depth (at least τ) w.r.t. the ranges in $\mathcal{S}$, ii) the ranges in $\mathcal{S}_i$ contribute at least δ fraction of the total weight of the ranges in $\mathcal{S}$ containing x and iii) fewer than $w_{\mathcal{S}_i}(x)$ ranges among the ranges in $\mathcal{S}_i$ containing x are sampled in the first stage.

Note that the second stage guarantees that the second condition in the lemma is satisfied for all elements in X. We now bound the probability that S_i is forced by x. Since each range $S \in \mathcal{S}$ is picked independently with probability $w(S)/(1-\delta)$ in the first stage, the expected value of s_i is $\mu = k_i(1-\delta)$ and by Chernoff bound, $\Pr(s_i < k_i) = \Pr(s_i < (1-\delta)\mu) \leq \exp(-\mu\delta^2/2) \leq \exp(-k_i\delta^2/2) \leq k_i^{-(c+5)}$ since $k_i \geq \delta k \geq \delta\tau = C\delta^{-2}\log\delta^{-1}$ which for large enough C implies that $k_i\delta^2/2 \geq (c+5)\log k_i$.

Thus, we have shown that the probability that the range S_i is forced by a particular element x having weighted depth k_i w.r.t. $\mathcal{S}_i$ is at most $k_i^{-(c+5)}$. However S_i may be forced by many elements in X. To bound the probability that S_i is forced (by some element), first note that all elements lying in the same cell of $(X, \mathcal{S}_i)$ behave identically i.e., they all either force S_i or don't force S_i. Thus, we only need to consider elements in distinct cells of $(X, \mathcal{S}_i)$. By Lemma 6, for any $t \geq 1$, S_i contains at most $O(t^{c+3})$ cells of depth t in $(X, \mathcal{S}_i)$. Since each range $S \in \mathcal{S}$ has a weight $w(S) \geq 0.5$, it follows that if an element of X has (unweighted) depth t, then its weighted depth is at least $t/2$. Thus, the probability that S_i is forced by elements in X of depth t w.r.t. $\mathcal{S}_i$ is at most $O(t^{c+3}) \cdot O(t^{-(c+5)}) = O(t^{-2})$. Since the second condition in the lemma is trivially satisfied for elements having weighted depth at most τ in $(X, \mathcal{S})$, we are concerned only with elements having weighted depth $k \geq \tau$ in $(X, \mathcal{S})$, and any such an element can force S_i only if its weighted depth in $(X, \mathcal{S}_i)$ is at least $\delta k \geq \delta\tau$, the probability that S_i is forced (by some element) is at most $\sum_{t=\delta\tau}^{\infty} O(t^{-2}) = O(1/(\delta\tau)) = O(\delta)$. The lemma follows. ◀

The following is an *unweighted* version of the above lemma.

▶ **Lemma 9** (Unweighted Sampling Procedure)**.** *Let $(X, \mathcal{S})$ be a c-linear set system and let $m = |\mathcal{S}|$. Let Δ be a large enough parameter. Then, there is a procedure to pick a subset $\mathcal{S}' \subseteq \mathcal{S}$ s.t.*

(i) $|\mathcal{S}'| \leq (1 + O(\sqrt[4]{(\log\Delta)/\Delta})\,|\mathcal{S}|/2$ *and*

(ii) *any element $x \in X$ having depth $k \geq \Delta$ in $\mathcal{S}$, has depth at least $k/2$ in $\mathcal{S}'$.*

Proof. The lemma follows from a straightforward application of Lemma 8 with parameter $\delta = \alpha\sqrt[4]{(\log\Delta)/\Delta}$ for a small enough constant α s.t. $\Delta \geq \frac{2}{\delta}C\,\delta^{-3}\log\delta^{-1}$ where C is the constant in Lemma 8 and setting the weight of each range in $\mathcal{S}$ to $0.5(1+\delta)$. ◀

We now restate and prove Theorem 2.

▶ **Theorem 2.** *Let $(X, \mathcal{S})$ be a c-linear set system for some constant c and let $m = |\mathcal{S}|$. Consider the set multicover problem defined by $(X, \mathcal{S})$ in which each element $x \in X$ has a demand $d(x)$. Let y be any feasible solution to the linear programming relaxation of this problem i.e., y satisfies the constraints: $\forall x \in X$, $\sum_{S \ni x} y_S \geq d(x)$, and let $\delta \in (0,1)$ be a given parameter. Then, we can obtain a subset of ranges $\mathcal{S}' \subseteq \mathcal{S}$ s.t.*

(i) $\forall x \in X$, $|\mathcal{S}'(x)| \geq d(x) - 3\tau$, *where* $\tau = C\delta^{-4} \log \delta^{-1}$ *with a large enough C*

(ii) $|\mathcal{S}'| \leq (1 + O(\delta)) \sum_{S \in \mathcal{S}} y_S$

Proof. Since condition (i) in the lemma is trivially satisfied for elements with demand at most 3τ, we can assume without loss of generality that each element has demand more than 3τ. We associate a *value* v_S with each range $S \in \mathcal{S}$. Initially $v_S = y_S$. The proof is constructive, and consists of two phases. The first phase consists of several rounds in which we modify the values of the ranges so that they are either 0 or lie in the interval $[0.5, 1]$. In a second phase we pick each range $S \in \mathcal{S}$ into $\mathcal{S}'$ with probability equal to the value v_S at the end of the first phase.

In the first phase we maintain a partition of the set of ranges $\mathcal{S}$ into two sets $\mathcal{F}$ and $\mathcal{N}$ where $\mathcal{F}$ is the set of *frozen* ranges whose values are either 0 or in the interval $[0.5, 1]$ and $\mathcal{N} = \mathcal{D} \setminus \mathcal{F}$ is the set of *non-frozen* ranges. We do not modify the value of any range in $\mathcal{F}$ which means that once a range is in $\mathcal{F}$, it remains in $\mathcal{F}$. We continue to modify the values of the ranges in $\mathcal{N}$ until all of them move to $\mathcal{F}$.

The first phase has several rounds. At the beginning of the first round, each range S has value y_S. The total value of all ranges at this time is $\sum_{S \in \mathcal{S}} y_S$. We replace each range $S \in \mathcal{N}$ by $\lfloor m \cdot y_S \rfloor$ replicas where each replica has value $\lambda_1 = 1/m$. The value of a range S is now the product of the number of replicas and the replica value. Note that the total value of any particular range decreases by at most $1/m$ due to this and therefore for each element $x \in X$, we have $\sum_{S \ni x : S \in \mathcal{S}} v_S \geq d(x) - 1$ i.e., v satisfies the demands with a *deficit* of ≤ 1.

At the beginning of round i, each replica has value $\lambda_i = 2^{i-1}/m$. Note that the replica values are uniform in any round and we double the replica value in each round. In round i, we use the unweighted sampling procedure (Lemma 9) with the parameter Δ set to $\Delta_i = \tau/\lambda_i$ to obtain a subset of the replicas that go to the round $i+1$. Note that the value of a frozen range is not modified by this procedure. The value of a non-frozen range changes according to the number of its replicas that make it to the next round. In particular, if a range has lost all its replicas, then its value becomes 0 and is frozen. Similarly, a range is frozen if the total value of its replicas becomes at least 0.5. The number of rounds is less than $\log_2 m$ since after so many rounds each replica has value at least 0.5.

Let v_S be the value of range S after the first phase. We claim that v satisfies $\forall x \in X$: $\sum_{S \in \mathcal{S}} v_S \geq d(x) - 2\tau$. To see this, consider any element $x \in X$ and let us say that x is *satisfied* at any point in time if the total value of the ranges containing it is at least its demand minus 1 i.e., $\sum_{S \ni x} v_S = \sum_{S \ni x : S \in \mathcal{N}} v_S + \sum_{S \ni x : S \in \mathcal{F}} v_S \geq d(x) - 1$. Note that x is satisfied at the beginning of round 1. Furthermore, if x is satisfied at the beginning of round i and in addition $\sum_{x \ni S : S \in \mathcal{N}} v_S \geq \tau$, then it is also satisfied after round i. This follows since every replica has value λ_i in round i, which means that if x's depth in the arrangement of the replicas is k then $k \geq \Delta_i = \tau/\lambda_i$ and Lemma 9 guarantees that its depth with respect to the replicas after round i is at least $k/2$. Recall that replica weights are doubled in every round which compensates for the decrease in the number of replicas covering x. Thus, if x is satisfied at the beginning of round i, then the only way it can be dissatisfied at the end of the round is if at the beginning of the round, $\sum_{S \ni x : S \in \mathcal{N}} v_S < \tau$ which means that $\sum_{S \ni x : S \in \mathcal{F}} v_S \geq d(x) - \tau - 1 \geq d(x) - 2\tau$. Since we don't modify the value of any range once it enters the set $\mathcal{F}$, this inequality is satisfied at the end of the first phase too.

We now argue that at the end of the first phase: $\sum_{S\in\mathcal{S}} v_S \le (1+O(\delta))\sum_{S\in\mathcal{S}} y_S$. To see this note that, by Lemma 9, the total value of the ranges in $\mathcal{S}$ increases, by a factor of at most $1+O(\sqrt[4]{(\log\Delta_i)/\Delta_i})$ in round i. Therefore, the total value of all ranges increases in the first phase by a factor of at most

$$\prod_{i=1}^{t}\left(1+O\left(\sqrt[4]{\frac{\log\Delta_i}{\Delta_i}}\right)\right) \le \exp\left[O\left(\sum_{i=1}^{t}\sqrt[4]{\frac{\log\Delta_i}{\Delta_i}}\right)\right] \le \exp\left[O\left(\sqrt[4]{\frac{\log\Delta_t}{\Delta_t}}\right)\right] \le (1+O(\delta))$$

where $t < \log_2 m$ is the number of rounds. The second last inequality above follows from the fact that $\sqrt[4]{(\log\Delta_i)/\Delta_i}$ increases geometrically with i since $\Delta_i = \tau m/2^{i-1}$ decreases geometrically with i. The last inequality follows from the fact that $t < \log_2 m$, which means that $\Delta_t < \tau$.

Recall that at the end of the first phase, the value of each range is either 0 or lies in the range $[0.5, 1]$. In the second phase, we apply Lemma 8 to the ranges having a non-zero value $v_S \in [0.5, 1]$ at the end of the first phase with the weight function $w(S) = v_S$ and the parameter δ. Since every element $x \in X$ is assumed to have demand at least 3τ, its weighted depth k with respect to the weight function w is least $d(x)-2\tau \ge \tau$. Lemma 8 then guarantees that its unweighted depth with respect to the set of ranges $\mathcal{S}'$ returned by Lemma 8 is also at least k. Thus $\mathcal{S}'$ satisfies the first condition in the lemma. It also satisfies the second condition in the lemma since Lemma 8 guarantees that $|\mathcal{S}'| \le (1+O(\delta))\sum_{S\in\mathcal{S}} v_S$ which is at most $(1+O(\delta))\sum_{S\in\mathcal{S}} y_S$ since we had established earlier that $\sum_{S\in\mathcal{S}} v_S \le (1+O(\delta))\sum_{S\in\mathcal{S}} y_S$. ◀

5 Set Multicover with bounded demands

Consider the set multicover problem defined by a set system $(X, \mathcal{S})$ and a demand function d s.t. for each element $x \in X$, $d(x)$ is bounded above by a constant Θ. We will show that a standard local search algorithm (see e.g. [2, 9, 28, 30] for details) yields a PTAS for this problem when X is a set of points and $\mathcal{S}$ is a set of non-piercing regions in the plane, or when X is a set of pseudodisks and $\mathcal{S}$ is a set of points in the plane. The algorithm takes as input a parameter k and starting with any feasible solution it tries to improve the solution by making *swaps* of size at most k and stops when no such improvement is possible.

In order to show that such a local search algorithm yields a PTAS, we consider an optimal solution R to the problem and a solution B returned by the local search algorithm. We want to show that $|B| \le (1+\epsilon(k))\,|R|$ where $\epsilon(k) \to 0$ as $k \to \infty$. We can assume without loss of generality that R and B are disjoint by removing the set $S = R \cap B$ from both and considering the problem with residual demands where the residual demand of any element $x \in X$ is $d(x) - |S(x)|$. Note that the residual demands are still bounded above by Θ and if $|B \setminus S| \le (1+\epsilon(k))\,|R \setminus S|$ then it follows that $|B| \le (1+\epsilon(k))\,|R|$.

By standard arguments [2, 9, 28, 30], it then suffices to show that there exists a graph $G = (R \cup B, E)$ satisfying the following ***local search conditions***:

1. For any $B' \subseteq B$, $(B \setminus B') \cup N(B')$ is a feasible solution, where $N(B')$ is the set of neighbors of B' in the graph G. We call this the *local exchange property.*
2. The graph has the *sublinear separator property* (see Definition 10).

▶ **Definition 10** (Sublinear Separator Property)**.** *We say that a graph H satisfies the sublinear separator property if there exist $0 < \alpha, \delta < 1$ s.t. for any induced subgraph H' of H with vertex set $V(H')$, there is a vertex separator $S \subset V(H')$ so that $|S| = O(|V(H')|^{\delta})$, and each connected component of $H' \setminus S$ is of size at most $\alpha|V(H')|$.*

In all applications of this technique so far, the critical part has been showing the existence of such a graph, which is what we now focus on. Instead of the local exchange property, we will use the following *local expansion* property which implies the local exchange property.

▶ **Definition 11** (Local Expansion). *Let R and B be two feasible solutions for a set multicover problem defined by a set system $(X, \mathcal{S})$ and a demand function d. We say that a graph $G = (R \cup B, E)$ satisfies the local expansion property with respect to $(X, R \cup B)$ if for every $x \in X$ at least one of the following statements hold:*

- *There are at least $d(x)$ elements in $R(x)$ that have $d(x)$ or more neighbors in $B(x)$.*
- *There are at least $d(x)$ elements in $B(x)$ that have $d(x)$ or more neighbors in $R(x)$.*

Here, by "neighbors" we mean neighbors in the graph G.

▶ **Proposition 12.** *Let R and B be two feasible solutions to the set multicover problem defined by a set system $(X, \mathcal{S})$ and a demand function d. Then, any graph $G = (R \cup B, E)$ satisfying the local expansion property with respect to $(X, R \cup B)$ also satisfies the local exchange property (i.e., local search condition 1).*

Proof. Consider any $x \in X$. The local expansion property implies that the subgraph of G induced on $R(x) \cup B(x)$ contains a matching of size $d(x)$. Let $U \subseteq R(x)$ and $V \subseteq B(x)$ be the matched elements. Suppose that we replace B by $(B \setminus B') \cup N(B')$ for some $B' \subseteq B$. Then we may lose $k \leq d(x)$ of elements of V but we gain at least an equal number of elements from U due to the matching. This implies that $(B \setminus B') \cup N(B')$ still satisfies x's demand. Since this is true for any $x \in X$, $(B \setminus B') \cup N(B')$ is a feasible solution. ◀

Observe that the intersection graph $I(R \cup B)$ of the ranges in $R \cup B$ in which two ranges are adjacent iff they share an element of X, has the local expansion property and therefore satisfies the local exchange property. However, the intersection graph may not satisfy the sublinear separator property. The next two lemmas show that for the geometric sets systems we consider, the intersection graph does satisfy the sublinear separator property if the depth of each cell in the set system is bounded above by a constant.

▶ **Lemma 13.** *Let $\mathcal{R}$ be a family of n non-piercing regions. Then, the number of pairs of regions in $\mathcal{R}$ that intersect at a point of depth 2 in the arrangement $\mathcal{A}(\mathcal{R})$ of $\mathcal{R}$ is at most $3n - 6$.*

Proof. Let $G = (\mathcal{R}, E)$ be the planar graph obtained from part 1 of Theorem 20. There must be an edge in G between any two regions in $\mathcal{R}$ that intersect at a point of depth 2 in $\mathcal{A}(\mathcal{R})$. The lemma follows from the fact that a planar graph with n vertices has at most $3n - 6$ edges. ◀

▶ **Lemma 14.** *Let $\mathcal{R}$ be a family of n non-piercing regions whose arrangement has depth at most Δ. Then, there exists a set Q of $O(\Delta n)$ points in the plane s.t. any intersecting pair of regions in $\mathcal{R}$ also intersect at one the points in Q.*

Proof. We assume that $\Delta \geq 2$ as otherwise no pair of regions in $\mathcal{R}$ intersect. Let Q be a minimal set of points so that any pair of intersecting regions in $\mathcal{R}$ also intersect at some point in Q. Since Q is minimal, for each $q \in Q$, there must be two regions A_q and B_q in $\mathcal{R}$ which intersect at q but don't intersect at any other point in Q. Suppose now that we take a sample $\mathcal{R}'$ of $\mathcal{R}$ by picking each region in $\mathcal{R}$ independently with probability $p = 1/\Delta$. Since any point $q \in Q$ is contained in at most Δ regions in $\mathcal{R}$, the probability that we pick A_q and B_q in our sample but do not pick any other regions containing q is at least $p^2(1-p)^{\Delta-2} \geq 1/(e\Delta)^2$. The expected number of points in Q for which this happens is at least $|Q|/(e\Delta)^2$. Each such point q corresponds to a distinct pair of regions in $\mathcal{R}'$ intersecting at a point of depth 2 (in the arrangement of $\mathcal{R}'$). Since the expected number of regions in $\mathcal{R}'$ is $pn = n/\Delta$, and there can be at most $3|\mathcal{R}'|$ pairs of regions in $\mathcal{R}'$ that intersect at a point of depth 2, by Lemma 13, we conclude that $|Q|/(e\Delta)^2 \leq 3n/\Delta$ which implies that $|Q| \leq 3e^2\Delta n$. ◀

▶ **Lemma 15.** *Let $H = (V, E)$ be a planar graph and let $\mathcal{T}$ be a set of subsets of V s.t. i) for each $T \in \mathcal{T}$ the subgraph $H[T]$ of H induced by T is connected and ii) each vertex $v \in V$ is contained in at most Δ of the sets in $\mathcal{T}$, where Δ is a constant. Then, the intersection graph of the sets in any $\mathcal{S} \subseteq \mathcal{T}$ has a balanced separator of size $O(\Delta\sqrt{|W|})$, where $W = \cup_{S\in\mathcal{S}} S$. In other words, the intersection graph of the sets in $\mathcal{T}$ satisfies the sublinear separator property.*

Proof. Fix any $\mathcal{S} \subseteq \mathcal{T}$. We assign a weight $w(v)$ to each vertex $v \in W$ as follows: each set $S \in \mathcal{S}$ has a weight of 1 and it distributes it equally among all the vertices it contains. More precisely, $w(v) = \sum_{S\in\mathcal{S}: v\in S} 1/|S|$. Note that the total weight of all vertices is $|\mathcal{S}|$. Let $X \subseteq W$ be a balanced vertex separator of $G = H[W]$ of size $O(\sqrt{|W|})$ so that the total weight of the vertices in each connected component of $G[W \setminus X]$ is at most $\alpha\,|\mathcal{S}|$ for some $\alpha < 1$. The fact that such a separator exists follows from the fact that $H[W]$ is planar and the planar graph separator theorem [27].

Let $\mathcal{X} = \{S \in \mathcal{S} : S \cap X \neq \emptyset\}$. As each vertex in X is contained in at most Δ sets in $\mathcal{S}$, $|\mathcal{X}| \leq \Delta|X| = O(\Delta\sqrt{|W|})$. Since each set $S \in \mathcal{S}$ induces a connected subgraph of H, $\mathcal{X}$ is a separator for the intersection graph I of the sets in $\mathcal{S}$. Furthermore, it is balanced since each connected component of $I[\mathcal{S} \setminus \mathcal{X}]$ corresponds to a connected component of $H[W \setminus X]$ and since the total weight of all vertices in the latter is at most αn, the number of sets in the former is also at most αn. ◀

▶ **Lemma 16.** *Let $\mathcal{R}$ be a family of n non-piercing regions in the plane whose arrangement has depth at most Δ, where Δ is a constant. Then the intersection graph of the regions in $\mathcal{R}$ has the sublinear separator property.*

Proof. Let Q be a set of $O(\Delta n)$ points obtained by applying Lemma 14. By Theorem 20, there is a planar graph $H = (Q, F)$, such that for any region $R \in \mathcal{R}$, the induced subgraph $H[Q_R]$, where $Q_R = R \cap Q$, is connected. Since every pair of intersecting regions in $\mathcal{R}$ intersect at some point in Q, the intersection graph of the sets in $\mathcal{S} = \{Q_R : R \in \mathcal{R}\}$ is isomorphic to the intersection graph of the regions in $\mathcal{R}$. Further, since the arrangement of the regions has depth Δ, no vertex of H is contained in more than Δ sets in $\mathcal{S}$. Therefore, the Lemma follows from the application of Lemma 15 to H and $\mathcal{S}$. ◀

▶ **Lemma 17.** *Let $(\mathcal{D}, P)$ be a set system defined by a set P of n points and a set $\mathcal{D}$ of pseudodisks in the plane. Then, the number of pairs of points (p, q) that lie in a common cell of depth 2 in $(\mathcal{D}, P)$ (i.e., there is a pseudodisk $D \in \mathcal{D}$ s.t. $D = \{p, q\}$) is at most $3n - 6$.*

Proof. The proof is the same as the proof of Lemma 13 except that we use part 2 of Theorem 20 instead of part 1. ◀

▶ **Lemma 18.** *Let P be a set of n points and let $\mathcal{D}$ be a family of pseudodisks in the plane s.t. for each $D \in \mathcal{D}, |D \cap P| \leq \Delta$. Then, there exists a subset $\mathcal{D}' \subseteq \mathcal{D}$ of size $O(\Delta n)$ s.t. any pair of points in P that belong to a common pseudodisk in $D \in \mathcal{D}$ also belong to a common pseudodisk $D' \in \mathcal{D}'$.*

Proof. The proof is identical to the proof of Lemma 14 except that we use Lemma 17 instead of Lemma 13. ◀

▶ **Lemma 19.** *Let P be a set of points and let $\mathcal{D}$ be a family of pseudodisks in the plane such that for any $D \in \mathcal{D}$, $|D \cap P| \leq \Delta$ for some constant Δ. Then, the graph $G(P, E)$ in which two points $p, q \in P$ are adjacent iff there is a pseudodisk $D \in \mathcal{D}$ containing both p and q, has the sublinear separator property.*

Proof. The proof is identical to the proof of Lemma 16, except that we use Lemma 18 instead of Lemma 14. ◀

In order to use the above lemmas, we need to modify a given set system $(X, R \cup B)$ so that each cell in it has bounded depth. We also require the following result which follows from Theorem 1 in [30].

▶ **Theorem 20** ([30]). *Let $\mathcal{R}$ be any family of non-piercing regions in the plane, and let P be any finite set of points in the plane. Then,*

1. *There exists a planar graph $G = (\mathcal{R}, E)$ s.t. for each point $p \in \mathbb{R}^2$, the subgraph of G induced by the regions containing p is connected.*
2. *There exists a planar graph $H = (P, E')$ s.t. for each $\gamma \in \mathcal{R}$, the subgraph of H induced by $\gamma \cap P$ is connected.*

5.1 Points and Non-Piercing regions

In this subsection, we prove Theorem 3. Consider the set multicover problem defined by a set system $(P, \mathcal{R})$ where P is a set of points and $\mathcal{R}$ is a family of non-piercing regions in the plane, and in which the demand $d(p)$ of each point is at most Θ. As before, $R, B \subseteq \mathcal{R}$ represent an optimal solution and a solution returned by the local search algorithm respectively. Furthermore, as argue before, we can assume without loss of generality that $R \cap B = \emptyset$. We will show the existence of a graph $G = (R \cup B, E)$ satisfying the local search conditions, implying a PTAS for this problem.

We first need a technical tool from [30] for modifying the regions so that the arrangement of the modified regions has bounded depth. A *maximal cell* in an arrangement of regions in the plane is a cell whose depth is higher than all neighboring cells.

▶ **Definition 21** (Cell Bypassing [30]). *Let $\mathcal{R}$ be a non-piercing family of regions. Let $\gamma \in \mathcal{R}$ be one of the regions and let C be a maximal cell contained in γ so that that the boundary of γ contributes exactly one arc to the boundary of C. Then, if we modified γ to $\gamma' = \gamma \setminus (C \oplus K_\epsilon)$, the resulting set of regions remains a non-piercing family. Here $\oplus$ denotes Minkowski sum and K_ϵ is a disk of arbitrarily small radius ϵ.*

We refer to the modification of γ to γ' in the above theorem as the "bypassing of C by γ". Note that the modified region $\gamma' \approx \gamma \setminus C$ since ϵ is arbitrarily small.

▶ **Lemma 22.** *There exists a graph $G = (R \cup B, E)$ that satisfies the local search conditions with respect to the set system $(P, \mathcal{R})$.*

Proof. We will construct a graph that satisfies the local expansion property at every point $p \in P$ and has the sublinear separator property. In fact, we will make the first condition stronger: we require the local expansion property to be satisfied at every point in the plane. In other words, we replace P by $\mathbb{R}^2$ and for each point $q \in \mathbb{R}^2$, we define the demand $d(q)$ as $\min(|R(q)|, |B(q)|, \Theta)$. Note that R and B are still feasible solutions to this modified problem and a graph $G = (R \cup B, E)$ satisfying the local expansion property with respect to $(\mathbb{R}^2, R \cup B)$ also satisfies the property with respect to $(P, R \cup B)$.

If the depth of the arrangement of $R \cup B$ is at most 2Θ, then we can simply use Lemma 16 to obtain the required graph G. Otherwise, consider a cell C of maximum depth. By the results in [30], there exists a region $\gamma \in R \cup B$ that contributes exactly one arc to the boundary of C. Our plan is to modify γ so that it bypasses C in order to reduce the depth of the cell C.

For convenience, we will refer to the regions in R and B as *red* and *blue* respectively. We will use the following slightly modified version of the cell bypassing procedure: The process remains the same as before except when the cell C to be bypassed consists of just one side i.e., it is identical to one of the regions. In such a case, the usual cell bypassing procedure removes the region. Instead, we do the following: if the points in the region are contained in more than Θ regions of its color, then we remove the region as before. Otherwise, if the point is contained in exactly Θ regions of its color (including itself), we keep the region but make it *inactive* so that this region does not participate in any further cell bypassing. Initially, all regions are defined to be active. Note that since an inactive region lies in the interior of a cell in the arrangement of the remaining active regions, it does not affect further cell bypassing. Only a cell defined by active regions is bypassed. All inactive regions lie in the interior of some cell in the arrangement of the active regions.

Let C be a cell of maximum depth in the arrangement of the active regions in $R \cup B$ and suppose that this cell is contained in at least $\Theta + 1$ active regions of some color (either red or blue). Let us assume that it is contained in at least $\Theta + 1$ active blue regions. The other case is analogous. We will argue that a graph satisfying the local expansion property for the modified red and blue regions obtained after bypassing C by one of the active regions also satisfies the condition for the original regions. To see this, consider any point p in the cell C that we are about to bypass in the arrangement of the active regions. The point p may be contained in some of the inactive regions contained in C apart from the active regions containing C. However, p cannot be contained in any inactive blue region β since that would mean that p was contained in at most Θ blue regions when β became inactive and is now contained in more than Θ blue regions (by assumption). Since cell bypassing only contracts regions, this is impossible.

Let γ be the region bypassing C. If γ is blue, note that p is still contained in at least $\Theta \geq d(p)$ blue regions after bypassing. Therefore, a graph satisfying the local expansion property for p in the arrangement of the modified regions also satisfies the condition in the arrangement of the original regions.

Consider now the case when γ is red. Let C' be a cell contained in γ and adjacent to C in the original arrangement of the active regions. Let q be a point in C' that is not contained in any of the inactive regions. Note that the set of red regions that contain at least one of the points in $\{p, q\}$ after bypassing is the same as the set of regions containing p before bypassing. Furthermore, both p and q are contained in the same set of blue regions before and after bypassing. Thus, a graph satisfying the locality preserving condition for the points p and q in the modified arrangement also satisfies the condition for p in the original arrangement.

Each cell-bypassing operation either decreases the maximum depth of the arrangement of the active regions, or decreases the number of cells of maximum depth in that arrangement. Therefore, we eventually obtain an arrangement where the maximum number of blue or red regions containing any cell in the arrangement of active regions is at most Θ. At this point, the number of active or inactive regions of the same color containing any point in the plane is at most Θ. Thus the depth of the arrangement of all active and inactive regions is at most 2Θ. We can now obtain the required graph using Lemma 16. ◀

Theorem 3 now follows from the discussions above.

5.2 Pseudodisks and Points

In this subsection, we prove Theorem 5. Consider a set multicover problem defined by the set system $(\mathcal{D}, P)$ where P is set of points and $\mathcal{D}$ is a family of non-piercing regions in the plane, and in which each pseudodisk $D \in \mathcal{D}$ has a demand bounded above by Θ. As earlier, let $R, B \subseteq P$ represent an optimal solution and a solution yielded by the local search algorithm respectively. Also, without loss of generality, $R \cap B = \emptyset$.

▶ **Lemma 23.** *There exists a graph $G = (R \cup B, E)$ that satisfies the local search conditions with respect to the set system $(\mathcal{D}, P)$.*

In order to prove Lemma 23, we will use the following result of Pinchasi [29].

▶ **Lemma 24** ([29]). *Let $\mathcal{D}$ be a set of pseudodisks in the plane. For any specified $D \in \mathcal{D}$ and any point $q \in D$, we can continuously shrink D to q so that the arrangement at any time remains an arrangement of pseudodisks.*

Proof of Lemma 23. If every pseudodisk in $\mathcal{D}$ contains at most 2Θ points of P, then we can use Lemma 19 to obtain the required graph. Now, consider any pseudodisk $D \in \mathcal{D}$ containing more than 2Θ points of P. Then, $\mathcal{D}$ must contain either $> d(D)$ points of B or $> d(D)$ points of R. Using Lemma 24, we first shrink D to D' as follows: we pick any point $p \in D \cap (R \cup B)$ and imagine continuously shrinking D to the point p. During the shrinking, we stop as soon as $|D \cap R| = d(D)$ or $|D \cap B| = d(D)$. We call this modified region D'. Lemma 24 guarantees that the arrangement obtained by replacing D by D' is still a family of pseudodisks.

Note that both $|D' \cap B|$ and $|D' \cap R|$ are at least $d(D)$ and one of them is equal to $d(D)$. Assume that $|D' \cap B| = d(D)$, the other case being analogous. Then, one by one, for every $b \in D' \cap B$, we imagine shrinking a copy of D' continuously to the point b and stop when the region contains exactly $d(D)$ points of R. We call this region D_b and add it to the collection $\mathcal{D}$. Again by Lemma 24 we can do this so that the arrangement remains an arrangement of pseudodisks. Finally, we remove the pseudodisk D from the collection $\mathcal{D}$. Observe that each of the regions added to the collection contain at most $d(D)$ points of R and at most $d(D)$ points of B.

This entire procedure is repeated for each pseudodisk $D \in \mathcal{D}$ containing more than 2Θ points. Let $\mathcal{D}'$ be the collection of regions obtained in the end. We now claim that a graph $G = (R \cup B, E)$ satisfying the local expansion property with respect to $(\mathcal{D}', R \cup B)$ also satisfies the local expansion property with respect to $(\mathcal{D}, R \cup B)$.

To see this, observe that if a pseudodisk $D \in \mathcal{D}$ is still in $\mathcal{D}'$, then a graph that satisfies the local expansion property with respect to $(\mathcal{D}, R \cup B)$ also satisfies the condition for D. On the other hand, if D was removed i.e., $D \notin \mathcal{D}'$, then one of the following holds:

- There are at least $d(D)$ points $b \in B \cap D$ for each of which we have added a pseudodisk D_b contained in D s.t. D_b contains $|d(D)|$ points of $R \cap D$
- There are at least $d(D)$ points $r \in R \cap D$ for each of which we have added a pseudodisk D_r contained in D s.t. D_r contains $|d(D)|$ points of $B \cap D$.

Assume the first case holds for a pseudodisk D removed from $\mathcal{D}$. The second case is analogous. Then, since each region D_b added has exactly $d(D)$ points of R, in any graph G that satisfies the local expansion property with respect to $(\mathcal{D}', R \cup B)$, the point b must be adjacent to all of the points in $R \cap D_b \subseteq D$. Since we have added such a pseudodisk D_b for at least $d(D)$ points in $B \cap D$, we have at least $d(D)$ points $B \cap D$ each of which has at least

$d(D)$ neighbors in $R \cap D$. Thus, G satisfies the local expansion property with respect to $(\mathcal{D}, R \cup B)$. Now, since each $D \in \mathcal{D}'$ has at most 2Θ points of $R \cup B$, Lemma 19 yields the required graph G that satisfies the local search conditions with respect to $(\mathcal{D}, R \cup B)$. ◀

Theorem 5 follows from the discussions above.

6 $(2+\epsilon)$-Approximation Algorithms

In this section, we prove Theorems 1 and 4, which require the following lemmas,

▶ **Lemma 25.** *The set system $(P, \mathcal{R})$ where P is a set of points and $\mathcal{R}$ is a non-piercing family is 2-linear.*

Proof. By Theorem 20 of [30], there is a planar support for the dual set system $(\mathcal{R}, P)$ i.e., there is a planar graph $H = (\mathcal{R}, E)$ s.t. for any $p \in P$, the set of regions $\mathcal{R}(p)$ induce a connected subgraph of H. Since H has $O(m)$ edges where $m = |\mathcal{R}|$, the number of cells of depth 2 in $(P, \mathcal{R})$ is $O(m)$. By a standard Clarkson-Shor type argument, this shows that the number of cells of depth at most k is $O(mk^2)$. ◀

Note that a family of non-piercing regions may not have linear union complexity but the set system $(P, \mathcal{R})$ still has a low shallow cell complexity.

▶ **Lemma 26.** *The set system $(\mathcal{D}, P)$ where $\mathcal{D}$ is a set of pseudodisks and P is a set of points in the plane is 2-linear.*

Proof. By Theorem 20 of [30], there is a planar support for the dual set system $(P, \mathcal{D})$ i.e., there is a planar graph $H = (P, E)$ s.t. for any $D \in \mathcal{D}$, the induced subgraph on the points in D is connected. Since H has $O(m)$ edges where $m = |P|$, the number of cells of depth 2 in $(P, \mathcal{D})$ is $O(m)$. By a standard Clarkson-Shor type argument, this shows that the number of cells of depth at most k is $O(mk^2)$. ◀

Proof of Theorem 1. We solve the LP-relaxation (1) for the multicover problem, and obtain a solution y, with objective value $\sum_{D \in \mathcal{R}} y_D \leq |\textsc{Opt}|$ where $\textsc{Opt}$ is an optimal solution to the given set multicover problem. By Lemma 25, the set system $(P, \mathcal{R})$ is 2-linear, and therefore we can use Theorem 2 to obtain a subset $\mathcal{R}' \subseteq \mathcal{R}$ of size at most $(1 + O(\delta)) \sum_{D \in \mathcal{R}} y_D \leq (1 + O(\delta)) |\textsc{Opt}|$, which satisfies the demands of the points in p with a deficit of at most $\Theta = O(\delta^{-4} \log \delta^{-1})$. We then consider the problem with the residual demands in which all the demands are at most Θ. For this problem, we obtain a solution $\mathcal{R}'' \subseteq \mathcal{R}$ of size at most $(1 + O(\delta)) |\textsc{Opt}|$ using Theorem 3. Then, $\mathcal{R}' \cup \mathcal{R}''$ is a solution to the set multicover and has size at most $(2 + O(\delta)) |\textsc{Opt}|$. By taking $\delta = \alpha\epsilon$ for some suitably small constant $\alpha > 0$, we get a $(2 + \epsilon)$-approximation to the optimal solution. ◀

Proof of Theorem 4. The proof is exactly the same as the proof of Theorem 1 except that we use Lemma 26 instead of Lemma 25 and Theorem 5 instead of Theorem 3. ◀

7 Conclusion

In this paper, we made progress on the Set Multicover in the geometric setting. We obtained a $(2 + \epsilon)$ approximation by combining LP-rounding and local search. We believe that local search itself yields a PTAS for the problem with arbitrary demands. Unfortunately we are currently able to prove that local search yields a PTAS only when the demands bounded above by a constant. Our approach does not work for the weighted setting since local search thus far works only for the unweighted setting. It is likely that a dynamic programming approach similar to [26, 16] can give a PTAS for the weighted set multicover problem for unit size disks and squares in the plane. Obtaining a PTAS for the problem with arbitrary size squares or disks seems challenging even for the set cover problem. Even with the improved approximation factor, our algorithm remains infeasible in practice due to the prohibitive running time ($n^{\text{poly}(1/\epsilon)}$) of the local search algorithm.

References

1. Pankaj K. Agarwal and Jiangwei Pan. Near-linear algorithms for geometric hitting sets and set covers. *Discret. Comput. Geom.*, 63(2):460–482, 2020. doi:10.1007/s00454-019-00099-6.
2. Rom Aschner, Matthew J. Katz, Gila Morgenstern, and Yelena Yuditsky. Approximation schemes for covering and packing. In *WALCOM: Algorithms and Computation*, volume 7748 of *Lecture Notes in Computer Science*, pages 89–100. Springer Berlin Heidelberg, 2013.
3. Nikhil Bansal and Kirk Pruhs. Weighted geometric set multi-cover via quasi-uniform sampling. *JoCG*, 7(1):221–236, 2016.
4. Aniket Basu Roy, Sathish Govindarajan, Rajiv Raman, and Saurabh Ray. Packing and covering with non-piercing regions. *Discrete & Computational Geometry*, 2018.
5. Hervé Brönnimann and Michael T. Goodrich. Almost optimal set covers in finite vc-dimension. *Discrete & Computational Geometry*, 14(4):463–479, 1995.
6. Norbert Bus, Shashwat Garg, Nabil H. Mustafa, and Saurabh Ray. Tighter estimates for ϵ-nets for disks. *Comput. Geom.*, 53:27–35, 2016. doi:10.1016/j.comgeo.2015.12.002.
7. Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 1576–1585, 2012.
8. Timothy M. Chan and Sariel Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. *Discrete & Computational Geometry*, 48(2):373–392, 2012.
9. Timothy M. Chan and Sariel Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. *Discrete & Computational Geometry*, 48(2):373–392, 2012. doi:10.1007/s00454-012-9417-5.
10. Timothy M. Chan and Qizheng He. Faster approximation algorithms for geometric set cover. *CoRR*, abs/2003.13420, 2020. arXiv:2003.13420.
11. Chandra Chekuri, Kenneth L. Clarkson, and Sariel Har-Peled. On the set multicover problem in geometric settings. *ACM Trans. Algorithms*, 9(1):9:1–9:17, 2012. doi:10.1145/2390176.2390185.
12. Chandra Chekuri, Sariel Har-Peled, and Kent Quanrud. Fast lp-based approximations for geometric packing and covering problems. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 1019–1038. SIAM, 2020. doi:10.1137/1.9781611975994.62.
13. Kenneth L. Clarkson and Kasturi R. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete & Computational Geometry*, 37(1):43–58, 2007. doi:10.1007/s00454-006-1273-8.
14. Vincent Cohen-Addad, Philip N. Klein, and Claire Mathieu. Local search yields approximation schemes for k-means and k-median in euclidean and minor-free metrics. *SIAM J. Comput.*, 48(2):644–667, 2019. doi:10.1137/17M112717X.
15. Vincent Cohen-Addad and Claire Mathieu. Effectiveness of local search for geometric optimization. In *Proceedings of the Thirty-first International Symposium on Computational Geometry*, SoCG '15, pages 329–343, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
16. Thomas Erlebach and Erik Jan van Leeuwen. PTAS for weighted set cover on unit squares. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings*, pages 166–177, 2010.

17 Guy Even, Dror Rawitz, and Shimon Shahar. Hitting sets when the vc-dimension is small. *Inf. Process. Lett.*, 95(2):358–362, 2005. doi:10.1016/j.ipl.2005.03.010.

18 Uriel Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, 1998. doi:10.1145/285055.285059.

19 Zachary Friggstad, Kamyar Khodamoradi, Mohsen Rezapour, and Mohammad R. Salavatipour. Approximation schemes for clustering with outliers. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 398–414, 2018.

20 Zachary Friggstad, Mohsen Rezapour, and Mohammad R. Salavatipour. Local search yields a PTAS for k-means in doubling metrics. *SIAM J. Comput.*, 48(2):452–480, 2019. doi:10.1137/17M1127181.

21 Matt Gibson and Imran A. Pirwani. Algorithms for dominating set in disk graphs: Breaking the $\log n$ barrier. In *Algorithms – ESA 2010 - 18th Annual European Symposium, Liverpool, United Kingdom, September 6–8, 2010, Proceedings*, pages 243–254, 2010.

22 David Haussler and Emo Welzl. epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987. doi:10.1007/BF02187876.

23 Bruno Jartoux and Nabil H. Mustafa. Optimality of Geometric Local Search. In *34th International Symposium on Computational Geometry (SoCG 2018)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 48:1–48:15, 2018.

24 János Komlós, János Pach, and Gerhard J. Woeginger. Almost tight bounds for epsilon-nets. *Discret. Comput. Geom.*, 7:163–173, 1992. doi:10.1007/BF02187833.

25 Erik Krohn, Matt Gibson, Gaurav Kanade, and Kasturi Varadarajan. Guarding terrains via local search. *Journal of Computational Geometry*, 5(1):168–178, 2014.

26 Jian Li and Yifei Jin. A PTAS for the weighted unit disk cover problem. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, pages 898–909, 2015.

27 Richard J. Lipton and Robert E. Tarjan. A separator theorem for planar graphs. *SIAM Journal on Applied Mathematics*, 36(2):177–189, 1979.

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

29 Rom Pinchasi. A finite family of pseudodiscs must include a “small” pseudodisc. *SIAM J. Discret. Math.*, 28(4):1930–1934, 2014. doi:10.1137/130949750.

30 Rajiv Raman and Saurabh Ray. Planar support for non-piercing regions and applications. In *26th Annual European Symposium on Algorithms, ESA 2018, August 20-22, 2018, Helsinki, Finland*, pages 69:1–69:14, 2018.

31 Kasturi Varadarajan. Weighted geometric set cover via quasi-uniform sampling. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, pages 641–648, 2010.

32 Vijay V. Vazirani. *Approximation algorithms.* Springer, 2001. URL: http://www.springer.com/computer/theoretical+computer+science/book/978-3-540-65367-7.