On the Set Multi-Cover Problem in Geometric Settings

Chandra Chekuri[*]   Kenneth L. Clarkson[†]   Sariel Har-Peled[‡]

March 6, 2009

Abstract

We consider the set multi-cover problem in geometric settings. Given a set of points P and a collection of geometric shapes (or sets) $\mathcal{F}$, we wish to find a minimum cardinality subset of $\mathcal{F}$ such that each point $p \in P$ is covered by (contained in) at least $d(p)$ sets. Here $d(p)$ is an integer demand (requirement) for p. When the demands $d(p) = 1$ for all p, this is the standard set cover problem. The set cover problem in geometric settings admits an approximation ratio that is better than that for the general version. In this paper, we show that similar improvements can be obtained for the multi-cover problem as well. In particular, we obtain an $O(\log \text{opt})$ approximation for set systems of bounded VC-dimension, and an $O(1)$ approximation for covering points by half-spaces in three dimensions and for some other classes of shapes.

1 Introduction

The ***set cover*** problem is the following. Given a universe U of n elements and a collection of sets $\mathcal{F} = \{S_1, \ldots, S_m\}$ where each S_i is a subset of U, find a minimum cardinality sub-collection $C \subseteq \mathcal{F}$ such that C covers U; in other words, the union of the sets in C is U. In the weighted version each set S_i has a non-negative weight w_i and the goal is to find a minimum weight cover C. In this paper, we are primarily interested in a generalization of the set cover problem; namely, the ***set multi-cover*** problem. In this version, each element $e \in \mathsf{U}$ has an integer demand or requirement $d(e)$ and a multi-cover is a sub-collection $C \subseteq \mathcal{F}$ such that for each $e \in \mathsf{U}$ there are $d(e)$ *distinct* sets in C that contain e.[1] The set cover problem and its variants arise directly and indirectly in a wide variety of settings and have numerous applications. Often $\mathcal{F}$ is available only in an implicit fashion and could be exponential in the size of U or even infinite (for example $\mathcal{F}$ could be the set of all disks in the plane). The set cover problem is NP-HARD and consequently approximation algorithms for it have received considerable attention. A simple greedy algorithm, that iteratively adds a set from $\mathcal{F}$ that covers the most uncovered elements, is known to give a $(1 + \ln n)$ approximation, where $n = |\mathsf{U}|$. (In the weighted case, the algorithm picks the set with minimum average cost for the uncovered elements.) Similar bounds can also be achieved via rounding a linear programming relaxation. The advantage of the greedy algorithm is that it is also applicable in settings where $\mathcal{F}$ is given implicitly, but there exists a polynomial time oracle to (approximately) implement the greedy step in each iteration. It is also known that unless $\mathrm{P} = \mathrm{NP}$ there is no $o(\log n)$

---

[1]A related and somewhat easier variant allows a set to be picked multiple times. In this paper, unless explicitly stated, we use multi-cover for the version where only one copy of a set is allowed to be picked.

1

approximation for the set cover problem [LY94]. Moreover, unless $\mathrm{NP} \subset \mathrm{DTIME}(n^{O(\log\log n)})$ there is no $(1-o(1))\ln n$ approximation [Fei98]. Thus the approximability of the general set cover problem is essentially resolved if $\mathrm{P} \neq \mathrm{NP}$. However, there are many set systems of interest for which the hardness of approximation result does not apply. There has been considerable effort to understand the approximability of set cover in restricted settings. Set systems that arise in geometric settings are the focus of this paper. Previous work has shown that the set cover problem admits improved approximation ratios in various geometric cases.

In the geometric setting, we use $(\mathsf{P}, \mathcal{F})$ to describe a set system (also referred to as a range space) where P is a set of points and $\mathcal{F}$ is a collection of sets (also called *objects* or *ranges*). We are typically interested in the case where $\mathcal{F}$ is a set of "well-behaved shapes". Examples of such shapes include disks, pseudo-disks, and convex polygons. The goal is to cover a given finite set of points P in $\mathbb{R}^d$ by a collection of objects from $\mathcal{F}$. At a higher level of abstraction, one can consider set systems of small (or constant) VC dimension. In addition to the inherent theoretical interest in geometric set systems, there is also motivation from applications in wireless and sensor networks. In these applications the coverage of a wireless or sensor transmitter can be reasonably approximated as a disk-like region in the plane. The problem of locating transmitters to optimize various metrics of coverage and connectivity is a well-studied topic; see [TWDJ08] for a survey.

Brönnimann and Goodrich [BG95], extending the work of Clarkson [Cla93], used the reweighting technique to give an $O(\log \mathrm{opt})$ approximation for the set cover problem when the VC dimension of the set system is bounded[2]. Here opt is the size of an optimum solution. Known hardness results [LY94] preclude such an approximation ratio for the general set cover problem. The reweighting technique and its application to set cover [Cla93, BG95] show that the approximation ratio for set cover can be related to bounds on ε-nets for set systems. Using this observation, [BG95] showed an improved $O(1)$ approximation ratio for the set cover problem in some cases, including the problem of covering points by disks in the plane. Long [Lon01] made an explicit connection between the integrality gap of the natural LP relaxation for the set cover problem and bounds on the ε-nets for the set system (see also [ERS05]). This allows opt in the approximation ratio to be replaced by f, where f is the value of an optimum solution to the LP relaxation (i.e., the optimal fractional solution). Clarkson and Varadarajan [CV07] developed a framework to obtain useful bounds on the ε-net size via bounds on the union complexity of a set of geometric shapes. Using this framework they gave improved approximations for various set systems/shapes. Recently, Aronov, Ezra and Sharir [AES09], and Varadarajan [Var09] slightly improve the bounds of Clarkson and Varadarajan [CV07].

The geometric set cover problem induced by covering points by disks in the plane is strongly NP-Hard [FG88]; existing results yield a constant factor approximation although a PTAS for this problem has so far not been ruled out. Some other geometric coverage problems are known to be APX-hard [FMZ07]; that is, there is a constant $c > 1$ such that unless $\mathrm{P} = \mathrm{NP}$, there is no c approximation for them.

Our results. In this paper, we consider the multi-cover problem in the geometric setting. In addition to the set system $(\mathsf{P}, \mathcal{F})$, each point $\mathsf{p} \in \mathsf{P}$ has an integer demand $\mathsf{d}(\mathsf{p})$. Now a cover needs to include, for each point p, $\mathsf{d}(\mathsf{p})$ sets that contain p. For general set systems, the greedy algorithm and other methods such as randomized rounding, which work for the set cover problem, can be adapted to the multi-cover problem, yielding a $(1+\ln n)$ approximation [Vaz01]. In contrast, the ε-net based approach for geometric set cover does not generalize to the multi-cover setting in

---

[2]Brönnimann and Goodrich [BG95] consider the *hitting set* problem which is the set cover problem in the dual range space. In this paper we blur the distinction between set cover and hitting set.

a straight-forward fashion. Nevertheless, we are able to use related ideas, in a somewhat more sophisticated way, to obtain approximation ratios for the geometric set multi-cover problem that essentially match the ratios known for the corresponding set cover problem. In particular, we obtain the following bounds. In all the bounds, $\mathsf{f} \leq \mathsf{opt}$ is the value of an optimum (fractional) solution to the natural LP relaxation, and opt is the value of an optimum (integral) solution.

- $O(\log \mathsf{f})$ approximation for set multi-cover of set systems of bounded VC dimension.

- $O(1)$ approximation for (multi) covering points in $\mathbb{R}^3$ by halfspaces. This immediately leads to a similar result for multi-cover of disks by points in the plane.

- $O(\log \log \log \mathsf{f})$ approximation for covering points by fat triangles (or other fat convex polygonal shapes of constant descriptive complexity) in the plane.

The second and third results follow from a general framework for a class of "well-behaved" shapes based on the union complexity of the shapes. This is inspired by a similar framework from [CV07, AES09]. Our work differs from previous work for set cover in geometric settings in two ways. First, we use the LP relaxation in an explicit fashion in several ways, demonstrating its effectiveness. Second, our work points out the usefulness of shallow cuttings for the multi-cover problem. We hope that these directions will be further developed in the future.

2 Preliminaries

2.1 Problem statement and notation

Let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be a given set system with VC dimension δ. Here P is a set of points, and $\mathcal{F}$ is a collection of subsets of P, called *ranges* or *objects*. Assume that every point $\mathsf{p} \in \mathsf{P}$ has an associated integral ***demand*** $\mathsf{d}_\mathcal{I}(\mathsf{p}) \geq 0$. When the relevant set system is understood, we may write $\mathsf{d}(\mathsf{p})$. Here we would like to find a minimum cardinality set of ranges of $\mathcal{F}$ that covers P, such that every $\mathsf{p} \in \mathsf{P}$ is covered at least $\mathsf{d}(\mathsf{p})$ times. Note that we allow a range of $\mathcal{F}$ to be included only once in the cover. This is an instance of the ***set multi-cover*** problem. There is also a weaker version of the problem, where the solution may be a multiset; that is, a range may be included multiple times.

We will also discuss the demand of a set $\mathsf{P}' \subset \mathsf{P}$, which is $\mathsf{d}(\mathsf{P}') = \mathsf{d}_\mathcal{I}(\mathsf{P}') = \sum_{\mathsf{p} \in \mathsf{P}'} \mathsf{d}(\mathsf{p})$. The ***total demand*** of a set system $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ is $\mathsf{d}(\mathsf{P})$.

Definition 2.1 For a point $\mathsf{p} \in \mathsf{P}$ and a set $X \subseteq \mathcal{F}$ where each range in $\mathcal{F}$ has a non-negative weight, let $\#(\mathsf{p} \cap X)$ denote the ***depth*** of p in X; namely, it is the total weight of the ranges of X covering p. If the ranges do not have weights then we treat them as having weight one.

Definition 2.2 Given a multiset $\mathcal{Z} \subseteq \mathcal{F}$, let $\mathcal{J} = (\mathsf{Q}, \mathcal{G}) = (\mathsf{P}, \mathcal{F}) \setminus \mathcal{Z}$ denote the ***residual*** set system. The residual instance encodes what remains to be covered after we use the coverage provided by $\mathcal{Z}$. Each $\mathsf{p} \in \mathsf{P}$ has ***residual demand*** $\mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{Z}) = \max(\mathsf{d}(\mathsf{p}) - \#(\mathsf{p} \cap \mathcal{Z}), 0)$, and Q comprises the points of P with nonzero residual demand. Thus $\mathsf{d}_\mathcal{J}(\mathsf{p}) = \mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{Z})$. Also $\mathcal{G} = \mathcal{F} \setminus \mathcal{Z}$. We will also write, for $\mathsf{Q}' \subset \mathsf{Q}$, $\mathsf{d}_{\mathrm{res}}(\mathsf{Q}', \mathcal{Z}) = \sum_{\mathsf{p} \in \mathsf{Q}'} \mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{Z})$. In particular, $\mathsf{d}_{\mathrm{res}}(\mathsf{Q}, \mathcal{Z}) = \mathsf{d}_\mathcal{J}(\mathsf{Q})$ is the ***total residual demand*** of $\mathcal{I}$, with respect to Q.

A set system $(\mathsf{P}, \mathcal{F})$ has VC dimension δ if no subset of P of cardinality greater than δ is *shattered* by $\mathcal{F}$. Here a set $\mathsf{P}' \subseteq \mathsf{P}$ is shattered if for every $X \subset \mathsf{P}'$ there is a range $r \in \mathcal{F}$ such that $X = r \cap \mathsf{P}'$. Given a range space $S = (\mathsf{P}, \mathcal{F})$ its dual set system is $S^* = (\mathcal{F}, \mathsf{P}^*)$ where $\mathsf{P}^* = \{\mathcal{F}_\mathsf{p} \mid \mathsf{p} \in \mathsf{P}\}$ and $\mathcal{F}_\mathsf{p} = \{r \in \mathcal{F} \mid \mathsf{p} \in r\}$. For a set system S with VC dimension δ, we denote by δ^* the VC dimension of

S^*. It is known that $\delta^* \leq 2^{\delta+1}$ [Har08]; thus if S has bounded VC dimension, so does S^*. However, for specific set systems of interest, in particular geometric set systems, one can directly show much stronger upper bounds on δ^*.

2.2 LP relaxation

A standard approach to computing an approximate solution to an NP-hard problem is to solve a linear programming relaxation (LP) of the problem and round its fractional solution to an integral solution to the original problem.

In our case, if $\mathcal{F} = \{r_1, \ldots, r_m\}$ and $\mathsf{P} = \{\mathsf{p}_1, \ldots, \mathsf{p}_n\}$. The natural LP has a variable x_i for range r_i:

$$
\begin{aligned}
\min \sum_{i=1}^m & x_i \\
\sum_{i:\mathsf{p}_j \in r_i} x_i &\geq \mathsf{d}(\mathsf{p}_j) & \forall \mathsf{p}_j \in \mathsf{P}, \\
x_i &\in [0,1] & i = 1, \ldots, m.
\end{aligned}
\tag{1}
$$

Note that LP is a relaxation of the integer program for the set multi-cover problem in which x_i are required to take a value in $\{0, 1\}$. If repetitions of a set are allowed then the constraint $x_i \in [0, 1]$ is replaced by $x_i \geq 0$.

Let $\mathsf{f} = \mathsf{f}(\mathcal{I})$ denote the ***value*** of an optimum solution to the above LP. Clearly, opt $\geq \mathsf{f}(\mathcal{I})$. We will refer to the values assigned to the variables x_i for some particular optimal solution to the LP as the ***fractional solution***. In the following, we will refer to the value of x_i in the solution as the ***weight*** of the range r_i. We will sometimes use vectors that are not optimal solutions for LP, but only ***feasible***; that is, they satisfy the constraints.

2.3 Overview of Rounding for Geometric Set Cover

We briefly explain the previous approaches for obtaining approximation algorithms for the set cover problem in geometric settings. The work of Clarkson [Cla93] and Brönnimann and Goodrich [BG95] used the reweighting technique and ε-nets to obtain algorithms that provide approximation bounds with respect to the integer optimum solution. In [Lon01, ERS05], it was pointed out that these results can be reinterpreted as rounding the LP relaxation and hence the approximation bounds can also be stated with respect to the fractional optimum solution. Here we discuss this interpretation.

Note that in the set cover setting $\mathsf{d}(\mathsf{p}) = 1$ for all points. Consider a fractional solution to the LP given by x_i assigned to ranges $r_i \in \mathcal{F}$, with total value $\mathsf{f} = \sum_i x_i$. Let $\varepsilon = 1/\mathsf{f}$. From the constraint (Eq. (1)) it follows that for each p, $\sum_{i:\mathsf{p}\in r_i} x_i/\mathsf{f} \geq \mathsf{d}(\mathsf{p})/\mathsf{f} = 1/\mathsf{f} = \varepsilon$. Interpreting x_i/f as the weight of range r_i, we obtain a set system in which all points are covered to within a weight of ε. Therefore an ε-net of the (weighted) *dual* range space is a set cover for the original instance. Now one can plug known results on the size of ε-nets for set systems to immediately derive an approximation. For example, set systems with VC dimension δ have ε-nets of size $O(\delta/\varepsilon \cdot \log 1/\varepsilon)$ [PA95] and hence one concludes that there is a set cover of size $O(\delta^* \mathsf{f} \log \mathsf{f})$ computable in polynomial time, that is, an $O(\delta^* \log \mathsf{f})$ approximation. For some set systems improved bounds on the ε-net size are known. For example, if P is a finite set of points and $\mathcal{F}$ is a set of disks in the plane then ε-nets of size $O(1/\varepsilon)$ are known to exist for the dual set system and hence one obtains an $O(1)$ approximation for covering points by disks in the plane. Clarkson and Varadarajan [CV07] showed that bounds on the size of ε-nets can be obtained in the geometric setting from bounds on the union complexity of

objects in $\mathcal{F}$. We remark that the connection to ε-nets above also holds in the converse direction: for a given set system, the integrality gap of LP can be used to obtain bounds on the ε-net size.

In the multi-cover setting we can take the same approach as above. However, now we have for a point p, $\sum_{i:\mathsf{p}\in r_i} x_i/\mathsf{f} \geq \mathsf{d}(\mathsf{p}) \cdot \varepsilon$ where $\varepsilon = 1/\mathsf{f}$. Note that we now have non-uniformity due to different demands and hence an ε-net would not yield a feasible multi-cover for the original problem.

3   Multi-cover in spaces with bounded VC dimension

In this section, we prove the following theorem.

Theorem 3.1 *Let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be an instance of multi-cover with* VC *dimension δ and let δ^* be the* VC *dimension of the dual set system. There is a randomized poly-time algorithm that on input $\mathcal{I}$ outputs $O(\delta^*\mathsf{f}\log\mathsf{f})$ sets of $\mathcal{F}$ that together satisfy $\mathcal{I}$, where f is the value of an optimum fractional solution to $\mathcal{I}$.*

We have an easy proof of the above theorem for the setting in which a set is allowed to be used multiple times; the proof is based on results on relative approximations. See Section 3.1 for details.

We believe that it may be possible to adapt the proof of the relative approximation result to prove the above theorem for the setting in which a set is not allowed to be included more than once. Instead, in this version of the paper, we give another proof that uses the LP to reduce the problem to a regular set cover problem with a modified set system whose VC dimension is at most $O(\delta)$.

Geometric intuition.   Imagine we have a set of points and a set of disks $\mathcal{F} = \{r_1, \ldots, r_m\}$ (i.e., the ranges) in the plane. We solve the LP for this system. This results in weight assigned to each disk, such that the total weight of the disks covering a point $\mathsf{p} \in \mathsf{P}$ exceeds (or meets) its demand $\mathsf{d}(\mathsf{p})$. We add another dimension (we are now in 3d), and for each $i = 1, \ldots, m$ translate the disk $r_i \in \mathcal{F}$ to the plane $z = i$. Let $\mathcal{F}'$ denote the resulting set of m 2d disks that "live" in 3d. Observe that the projection of $\mathcal{F}'$ to the xy plane is $\mathcal{F}$. Every point $\mathsf{p}_j \in \mathsf{P}$ is now a vertical line ℓ_j (parallel to the z-axis), and we are asking for a subset X of $\mathcal{F}'$, such that every line ℓ_j stabs at least $\mathsf{d}(\mathsf{p}_j)$ disks of X. The fractional solution for the original problem induces a fractional solution to the new problem. The next step, is to break every line ℓ_j into segments, such that the total weight of the disks of $\mathcal{F}'$ intersecting a vertical segment is at least 1 (and at most 2). Let L' be this resulting set of segments. Consider the "set system" $\mathsf{S} = (L', \mathcal{F}')$, and its associated set cover instance of the disks of $\mathcal{F}'$ so that they intersect all the segments of L'. It is easy to verify that any solution of this set cover problem, is in fact a solution to the original multi-cover problem, and vice versa (up to small constant multiplicative error, say 2). We know how to solve such set-cover problems using standard tools. The key observation is that the projection of $(L', \mathcal{F}')$ on to the plane yields the original range space. Similarly, projecting $(L', \mathcal{F}')$ on to the z-axis results in a range space where the points are on the real line and the ranges are intervals. Since the range space $(L', \mathcal{F}')$ is the intersection of two range spaces of low VC dimension, it has low VC dimension. This implies that the set-cover problem on $(L', \mathcal{F}')$ has a good approximation [BG95] and this leads to a good approximation to the original multi-cover problem on S.

More formal solution.   Consider a fractional solution x to the LP associated with $\mathcal{I}$. If any set $r_i \in \mathcal{F}$ satisfies $x_i \geq 1/4$ then we add r_i to our solution. There can be at most $4\sum_i x_i = 4\mathsf{f}$ such sets, so including them does not harm our goal of a solution with $O(\mathsf{f})$ sets. We now work with the residual instance and hence we can assume that the fractional solution has no set r_i with $x_i \geq 1/4$.

Now, assume that we have fixed the numbering of the ranges of $\mathcal{F} = \{r_1, \ldots, r_m\}$, and consider the fractional solution, with the value x_i associated with r_i, see Eq. (1). In particular, for a point $\mathsf{p} \in \mathsf{P}$, consider the linear inequality

$$\sum_{i : \mathsf{p} \in r_i} x_i \geq \mathsf{d}(\mathsf{p}) .$$

This inequality holds for the fractional solution. We split this inequality into $O(\mathsf{d}(\mathsf{p}))$ inequalities having 1 on the right hand side. To this end, scan this inequality from left to right, and collect enough terms on the left-hand side, such that their sum (in the fractional solution) is larger than $1/2$. We will write down the resulting inequality, and continue in this fashion till all the terms of this inequality are exhausted.

Formally, let $U_0 = U = \left\{ i \ \middle| \ \mathsf{p} \in r_i \right\}$ be the sequence of indices of the ranges participating in the above summation, where U and U_0 are sorted in increasing order. For $\ell \geq 1$, let V_ℓ be the shortest prefix of $U_{\ell-1}$ such that $\sum_{i \in V_\ell} x_i \geq 1/2$, and let u_ℓ be the largest number (i.e., index) in V_ℓ, and let $U_\ell = (U_{\ell-1} \setminus V_\ell)$. Since each $x_i < 1/4$ we have that $\sum_{i \in V_\ell} x_i < 1/2 + 1/4 < 3/4$. We stop when $\sum_{i \in U_\ell} x_i < 1/2$ for the first time. This process creates $h \geq \mathsf{d}(\mathsf{p})$ inequalities of the form

$$\sum_{i \in V_\ell} x_i \geq 1/2,$$

for $\ell = 1, \ldots, h$. We obtain at least $\mathsf{d}(\mathsf{p})$ inequalities from the fact that $\sum_{i : \mathsf{p} \in r_i} x_i \geq \mathsf{d}(\mathsf{p})$ and by our observation that $\sum_{i \in V_\ell} x_i < 3/4$.

We next describe a new set system $(\mathsf{P}', \widehat{\mathcal{F}})$, derived from this construction of inequalities, such that a set cover solution to the new system implies a multi-cover solution to the original system, and the new system has small VC dimension.

The new set system $(\mathsf{P}', \widehat{\mathcal{F}})$ is defined as follows. For a p which was processed as above we create h copies, one for each V_ℓ. Each such copy of p corresponds to an interval $I = [\alpha, \beta]$, where α is $\min_{i \in V_\ell} i$, and β is $\max_{i \in V_\ell} i$. So p has h such intervals associated with it, say $I_1, \ldots, I_h$. We generate h new pairs from p, namely, $Q(\mathsf{p}) = \{(\mathsf{p}, I_1), \ldots, (\mathsf{p}, I_h)\}$.

We set $\mathsf{P}' = \cup_\mathsf{p} Q(\mathsf{p})$, and $\widehat{\mathcal{F}} = \left\{ \widehat{r_i} \ \middle| \ r_i \in \mathcal{F} \right\}$, where

$$\widehat{r_i} = \left\{ (\mathsf{p}, I) \in \mathsf{P}' \ \middle| \ \mathsf{p} \in r_i \text{ and } i \in I \right\}. \tag{2}$$

Note that $|\widehat{r_i}| = |r_i|$, and it can be interpreted as deciding, for each point $\mathsf{p} \in r_i$, which one of its copies should be included in $\widehat{r_i}$.

The following two claims follow easily from the construction.

Claim 3.2 *For the set cover instance defined by $(\mathsf{P}', \widehat{\mathcal{F}})$ there is a fractional solution of value $2 \sum_i x_i \leq 2\mathsf{f}$.*

Claim 3.3 *An integral solution of value β to the set cover instance $(\mathsf{P}', \widehat{\mathcal{F}})$ implies a multi-cover to the original instance of cardinality at most β.*

We need the following easy lemma on the dimension of intersection of two range spaces with bounded VC dimension.

Lemma 3.4 ([Har08]) *Let $\mathsf{S} = (X, \mathcal{R})$ and $\mathsf{T} = (X, \mathcal{R}')$ be two range spaces of VC-dimension δ and δ', respectively, where $\delta, \delta' > 1$. Let $\widehat{\mathcal{R}} = \left\{ r \cap r' \ \middle| \ r \in \mathcal{R}, r' \in \mathcal{R}' \right\}$. Then, for the range space $\widehat{\mathsf{S}} = (X, \widehat{\mathcal{R}})$, we have that $\delta(\widehat{\mathsf{S}}) = O(\delta + \delta')$.*

6

The crucial lemma is the following.

Lemma 3.5 *The* VC *dimension of the set system* $(\mathsf{P}', \widehat{\mathcal{F}})$ *is* $O(\delta)$ *and the* VC *dimension of its dual set system is* $O(\delta^*)$.

Proof: We define two set systems $\left(\mathsf{P}', \widetilde{\mathcal{F}} \right)$ and $\left(\mathsf{P}', \overline{\mathcal{F}} \right)$ as follows. $\widetilde{\mathcal{F}} = \left\{ \widetilde{r_i} \ \middle| \ r_i \in \mathcal{F} \right\}$ where

$$\widetilde{r_i} = \left\{ (\mathsf{p}, I) \in \mathsf{P}' \ \middle| \ \mathsf{p} \in r_i \right\},$$

and $\overline{\mathcal{F}} = \left\{ \overline{r_i} \ \middle| \ r_i \in \mathcal{F} \right\}$, where $\overline{r_i} = \left\{ (\mathsf{p}, I) \in \mathsf{P}' \ \middle| \ i \in I \right\}$.

Note that $\widehat{r_i} = \widetilde{r_i} \cap \overline{r_i}$ (see Eq. (2)). Therefore $(\mathsf{P}', \widehat{\mathcal{F}})$ is formed by the intersection of ranges $(\mathsf{P}', \widetilde{\mathcal{F}})$ with ranges of $(\mathsf{P}', \overline{\mathcal{F}})$. Therefore the VC dimension of $(\mathsf{P}', \widehat{\mathcal{F}})$ is bounded by $O\left(\widetilde{\delta} + \overline{\delta} \right)$ where $\widetilde{\delta}$ and $\overline{\delta}$ are the VC dimensions of $(\mathsf{P}', \widetilde{\mathcal{F}})$ and $(\mathsf{P}', \overline{\mathcal{F}})$ respectively, by Lemma 3.4. We observe that the set system $(\mathsf{P}', \widetilde{\mathcal{F}})$ has the same VC dimension as that of $(\mathsf{P}, \mathcal{F})$ since we only duplicate points. The set system $(\mathsf{P}', \overline{\mathcal{F}})$ has constant VC dimension $\overline{\delta} = 3$ since it is the intersection system of points on the line with intervals.

For the second part of the claim we note that the dual set system of $(\mathsf{P}', \widehat{\mathcal{F}})$ is the intersection of the dual set system of $\left(\mathsf{P}', \widetilde{\mathcal{F}} \right)$ with VC dimension δ^* and the dual set system of $\left(\mathsf{P}', \overline{\mathcal{F}} \right)$ with VC dimension $O(1)$. We again apply Lemma 3.4. $\blacksquare$

Now we apply the known results on the integrality gap of the LP for set cover as discussed in Section 2.3. These results imply that for the set system $(\mathsf{P}', \widehat{\mathcal{F}})$ there is an integral set cover of value $O(\delta^* \mathsf{f} \log \mathsf{f})$ (here we use Claim 3.2 and Lemma 3.5). From Claim 3.3, there is a multi-cover for the original instance of the desired size. This completes the proof of the theorem. We observe that the algorithm is in fact quite simple. After solving the LP, pick each range r_i independently with probability $\min\{1, cx_i\}$ where $c = \alpha \cdot \delta^* \log \mathsf{f}$ for a sufficiently large constant α. With constant probability this yields a multi-cover.

3.1 Multi-cover in spaces with bounded VC dimension when allowing repetition

We consider the case where sets in $\mathcal{F}$ are allowed to be picked multiple times to cover a point. For this purpose we use relative approximations. The standard definition of relative approximation is the dual of what we give below.

Let $\alpha, \phi > 0$ be two fixed constants. For a set system $\mathcal{I} = (\mathsf{P}, \mathcal{F})$, recall from Definition 2.1 that $\#(\mathsf{p} \cap \mathcal{F})$ denotes the number of sets in $\mathcal{F}$ that contain the point p. A ***relative*** (α, ϕ)-***approximation*** is a subset $X \subseteq \mathcal{F}$ that satisfies

$$(1 - \alpha) \frac{\#(\mathsf{p} \cap \mathcal{F})}{|\mathcal{F}|} \leq \frac{\#(\mathsf{p} \cap X)}{|X|} \leq (1 + \alpha) \frac{\#(\mathsf{p} \cap \mathcal{F})}{|\mathcal{F}|}. \tag{3}$$

for each $\mathsf{p} \in \mathsf{P}$ with $\#(\mathsf{p} \cap \mathcal{F}) \geq \phi \cdot |\mathcal{F}|$. It is known [LLS01] that there exist subsets with this property of size $\frac{c\delta}{\alpha^2 \phi} \log \frac{1}{\phi}$, where c is an absolute constant, and δ is the VC dimension of the dual set system of $(\mathsf{P}, \mathcal{F})$. Indeed, any random sample of that many sets from $\mathcal{F}$ is a relative (α, ϕ)-approximation with constant probability. To guarantee success with probability at least $1 - q$, one needs to sample $\frac{c}{\alpha^2 \phi} \left(\delta \log \frac{1}{\phi} + \log \frac{1}{q} \right)$ elements of X, for a sufficiently large constant c [LLS01].

To apply relative approximation for our purposes we let N be a large integer such that Nx_i is an integer for each range r_i (since the x_i are rational such an N exists). We create a new set system $(\mathsf{P}, \mathcal{F}')$ where $\mathcal{F}'$ is obtained from $\mathcal{F}$ by duplicating each range $r_i \in \mathcal{F}$ Nx_i times. Thus $|\mathcal{F}'| = N\mathsf{f}$. From the feasibility of x for the LP we have that $\#(\mathsf{p} \cap \mathcal{F}') \geq N\mathsf{d}(\mathsf{p}) \geq N\mathsf{f}\mathsf{d}(\mathsf{p})/\mathsf{f}$ for each $\mathsf{p} \in \mathsf{P}$.

Now we apply the relative approximation result to $(\mathsf{P}, \mathcal{F}')$ with $\phi = 1/\mathsf{f}$ and $\alpha = 1/2$ to obtain a set $X \subset \mathcal{F}'$ such that $|X| = \Theta(\delta^* \mathsf{f} \log \mathsf{f})$ and with the property that for each $\mathsf{p} \in \mathsf{P}$,

$$\frac{\#(\mathsf{p} \cap \mathcal{F}')}{2\,|\mathcal{F}'|} \leq \frac{\#(\mathsf{p} \cap X)}{|X|}.$$

As such, we have

$$\#(\mathsf{p} \cap X) \geq \frac{|X|}{2} \cdot \frac{\#(\mathsf{p} \cap \mathcal{F}')}{|\mathcal{F}'|} \geq \frac{|X|}{2} \cdot \frac{N\mathsf{d}(\mathsf{p})}{N\mathsf{f}} = \mathsf{d}(\mathsf{p}) \cdot \Omega(\delta^* \log \mathsf{f}) \geq \mathsf{d}(\mathsf{p}),$$

as desired.

Note that X is picked from $\mathcal{F}'$ which has duplicate copies of sets from $\mathcal{F}$. We believe that the following rounding should yield a desired multi-cover without repetitions. First, pick each r_i into the multi-cover if $x_i \geq c\delta^*/\log \mathsf{f}$ for sufficiently large constant c. Then for the remaining ranges pick each one independently with probability $cx_i \cdot \delta^* \log \mathsf{f}$. Analyzing this would essentially require a careful walkthrough of the proof for relative approximations.

4 Multi-cover for Halfspaces in 3d and Generalizations

In this section, we show that improved approximations can be obtained for specific classes of set systems induced by geometric shapes of low complexity. In particular, we describe an $O(1)$ approximation for the multi-cover problem when the points are in $\mathbb{R}^3$, and the ranges are induced by halfspaces. The main idea of using cuttings extend also to other nice shapes. We outline the extensions and some applications in Section 5.

4.1 Total demand, Sampling, and Residual demand

We develop some basic ingredients that are useful in (randomly) rounding the LP solution. These ingredients apply to a generic multi-cover instance, not necessarily a geometric one, however we use the notation of points and ranges for the sake of continuity.

Lemma 4.1 *Given a multi-cover instance* $\mathcal{I} = (\mathsf{P}, \mathcal{F})$, *one can compute a cover for* $\mathcal{I}$ *of size no more than the total demand* $\mathsf{d}_\mathcal{I}(\mathsf{P})$.

Proof: Indeed, scan the (unsatisfied) points of P one by one, and pick, for each such point p, in an arbitrary fashion, $\mathsf{d}(\mathsf{p})$ ranges that cover it to be added to the solution. Clearly, the ranges that are picked satisfy all the demands and the number of ranges picked is at most $\sum_\mathsf{p} \mathsf{d}(\mathsf{p}) = \mathsf{d}_\mathcal{I}(\mathsf{P})$. ∎

Given an instance of multi-cover $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ and a feasible fractional solution x, a ***cx-sample*** for a scalar c is a random sample of $\mathcal{F}$, formed by independently picking each of the ranges $r_i \in \mathcal{F}$ with probability $\min\{1, cx_i\}$, where x_i is the value assigned to r_i by the fractional solution. (For the i with $cx_i \geq 1$, so that i is chosen with probability one, we will simply assume that such choices have been made, and the demand removed; that is, we assume that hereafter that $x_i \leq 1/c$. Since the number of such i is at most $c\mathsf{f}$, this step does not affect our goal of obtaining an output cover with $O(\mathsf{f})$ sets.)

8

Lemma 4.2 *Let $c \geq 4$ be a constant and let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be a multi-cover instance with an LP solution satisfying $x_i \leq 1/c$ for all i. If $\mathcal{R}$ is a cx-sample and $\mathsf{p} \in \mathsf{P}$ is a point with demand $\mathsf{d} = \mathsf{d}(\mathsf{p})$, then*

$$\mathbf{Pr}\Big[\mathsf{p} \text{ is not fully covered by } \mathcal{R}\Big] = \mathbf{Pr}\Big[\#(\mathsf{p} \cap \mathcal{R}) < \mathsf{d}\Big] \leq \exp\Big(-\frac{c}{4}\mathsf{d}\Big),$$

and $\mathbf{E}\Big[\mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{R})\Big] \leq \exp\Big(-\frac{c}{4}\mathsf{d}\Big)$.

Proof: Let X_i be the indicator variable which is equal to one if the cx-sample includes the range $r_i \in \mathcal{F}$, and is zero otherwise. Let $Y = \#(\mathsf{p} \cap \mathcal{R}) = \sum_{i:\mathsf{p} \in r_i} X_i$; observe that $\mu = \mathbf{E}[Y] \geq c\mathsf{d}$ using the facts that x is a feasible solution to LP, and $x_i \leq 1/c$ for all i. For $j \in [0, d]$, we apply the Chernoff inequality [MR95] and use the fact that $c \geq 4$ to obtain:

$$\begin{aligned}
\mathbf{Pr}\Big[\#(\mathsf{p} \cap \mathcal{R}) \leq \mathsf{d} - j\Big] &\leq \mathbf{Pr}\Big[Y < \mu(1 - (c-1)/c - j/\mu)\Big] \leq \exp\Big(-\frac{\mu}{2}\Big(\frac{c-1}{c} + \frac{j}{\mu}\Big)^2\Big) \\
&\leq \exp\Big(-\frac{\mu}{4} - \frac{3}{4}j\Big) \leq \exp\Big(-\frac{c}{4}\mathsf{d} - \frac{3}{4}j\Big).
\end{aligned}$$

The first statement of the lemma follows by substituting $j = 1$ and observing that the desired bound follows, and the second follows by using the fact that, for a random variable Z taking non-negative integral values, that $\mathbf{E}[Z] = \sum_{k>0} \mathbf{Pr}[Z \geq k]$. This implies

$$\begin{aligned}
\mathbf{E}\Big[\mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{R})\Big] &= \sum_{1 \leq j \leq \mathsf{d}} \mathbf{Pr}\Big[\#(\mathsf{p} \cap \mathcal{R}) \leq \mathsf{d} - j\Big] \leq \sum_{1 \leq j \leq \mathsf{d}} \exp\Big(-\frac{c}{4}\mathsf{d} - \frac{3}{4}j\Big) \\
&= \exp\Big(-\frac{c}{4}\mathsf{d}\Big) \sum_{1 \leq j \leq \mathsf{d}} \exp\Big(-\frac{3}{4}j\Big) \leq \exp\Big(-\frac{c}{4}\mathsf{d}\Big) \frac{1}{\exp(3/4) - 1} \leq \exp\Big(-\frac{c}{4}\mathsf{d}\Big),
\end{aligned}$$

as claimed. ∎

In the following, for $t \geq 1$, let

$$\mathsf{P}_t = \Big\{\mathsf{p} \in \mathsf{P} \,\Big|\, t \leq \mathsf{d}(\mathsf{p}) < 2t \,\Big\}.$$

The lemma below implies that if the number of points in the set system is "small" then the multi-cover problem can almost be solved in one round of sampling.

Lemma 4.3 *Suppose there is a probability distribution on a collection of multi-cover instances such that an instance $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ chosen from the distribution satisfies, for any $t \geq 1$,*

$$\mathbf{E}\Big[\, |\mathsf{P}_t| \,\Big] \leq V \cdot K^t,$$

where K and V are fixed parameters of the distribution. Then there is a value c depending on K, such that for any feasible fractional solution x to $\mathcal{I}$, a cx-sample $\mathcal{R}$ results in expected total residual demand $\mathsf{d}_{\mathrm{res}}(\mathsf{P}, \mathcal{R}) \leq V$; here the expectation is with respect to the randomness of $\mathcal{I}$ and the independent randomness of the cx-sample.

9

Proof: Let $\mathcal{R}$ be a cx-sample of $\mathcal{F}$ for fixed $c \geq 4 + 4\log K$. Let X be the subset of $\mathcal{R}$ with $x_i = 1$. Since $\mathcal{R}$ is also a cx-sample of $\mathcal{I} \setminus X$, we assume hereafter that X is empty; the result for general X follows by application of the result to $\mathcal{I} \setminus X$.

By applying Lemma 4.2 to the induced range space $(\mathsf{P}_t, \mathcal{F})$, we have

$$
\begin{aligned}
\mathbf{E}_{\mathcal{I},\mathcal{R}}\left[\mathsf{d}_{\mathrm{res}}(\mathsf{P}_t, \mathcal{R})\right] &\leq \mathbf{E}_{\mathcal{I}}\left[\sum_{\mathsf{p} \in \mathsf{P}_t} \mathbf{E}_{\mathcal{R}}\left[\mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{R})\right]\right] \leq \mathbf{E}_{\mathcal{I}}\left[|\mathsf{P}_t|\right]\exp\left(-\frac{c}{4}t\right) \leq VK^t\exp\left(-\frac{c}{4}t\right) \\
&\leq V\exp(-t(c/4 - \log K)) \leq V\exp(-t).
\end{aligned}
$$

Then, by linearity of expectation, we have

$$
\mathbf{E}\left[\mathsf{d}_{\mathrm{res}}(\mathsf{P}, \mathcal{R})\right] = \sum_{i=0}^{\infty}\mathbf{E}\left[\mathsf{d}_{\mathrm{res}}(\mathsf{P}_{2^i}, \mathcal{R})\right] \leq \sum_{i=0}^{\infty}V\exp\left(-2^i\right) \leq V.
$$

Thus, after cx-sampling, the residual instance has total expected demand bounded by V, as claimed. ∎

4.2 Clustering the given instance

The key observation is Lemma 4.3: it provides a sufficient condition for an $O(1)$ approximation. Of course, it might not be true (even in low dimensional geometric settings) that the number of points (i.e., the total residual demand) is small enough, as required to apply this lemma. We pre-process the given instance via an initial sampling step and then employ a clustering scheme that partitions the points into regions; we argue that these regions and an induced multi-cover instance on them satisfies the conditions of the lemma.

The ***depth*** of a simplex $\triangle$ in a set of weighted halfspaces is the minimum depth of any point inside $\triangle$, see Definition 2.1.

To perform the aforementioned clustering, we will use the shallow cutting lemma of Matoušek [Mat92]. We next state it in the form needed for our application, which is a special case of Theorem 5.1.

Lemma 4.4 *Given a set $\mathcal{F}$ of weighted halfspaces in $\mathbb{R}^3$, with total weight W, there is a randomized polynomial-time algorithm that generates a set Γ of simplices, called a $(1/4W)$-**cutting**, with the following properties: the union of the simplices covers $\mathbb{R}^3$; the total weight of the boundary planes of $\mathcal{F}$ intersecting any simplex of Γ is bounded by $1/4$; and finally, for any $t \geq 0$, the expected total number of simplices of depth at most t is $O(Wt^2)$. (Here the expectation is with respect to the randomness of the algorithm.)*

4.2.1 The algorithm

Given an instance of multi-cover $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ of points and halfspaces in $\mathbb{R}^3$, our algorithm first computes the fractional solution to the LP induced by $\mathcal{I}$, yielding weights x_i. Next, for β an absolute constant in $(0, 1/4)$ to be specified later, we put in set X all ranges r_i with $x_i \geq \beta$. Let $(\mathsf{P}', \mathcal{F}') = (\mathsf{P}, \mathcal{F}) \setminus X$. Let $\mathsf{f}' = \sum_{r_i \in \mathcal{F} \setminus X} x_i$ be the total weight of the remaining ranges.

The remainder of the algorithm uses a auxiliary abstract multi-cover instance derived using cuttings, as described next.

Using the weights x_i, we build a $(1/4\mathsf{f}')$-cutting Γ for $\mathcal{F}'$. This induces an abstract multi-cover instance $(\Gamma, \mathcal{F}')$, where a simplex $\triangle \in \Gamma$ is covered by halfspace $h \in \mathcal{F}'$ only if the interior of $\triangle$ does not meet the boundary plane of h. The demand $\mathsf{d}(\triangle)$ is defined to be $\max_{\mathsf{p} \in \mathsf{P} \cap \triangle}\mathsf{d}_{\mathrm{res}}(\mathsf{p}, \mathcal{F}')$.

The weights x_i give a feasible fractional solution to $\mathcal{I} \setminus X$, and so the depth of Δ is at least $\mathsf{d}(\Delta) - 1/4$, where Δ "loses" at most weight $1/4$ of depth due to halfspaces whose boundary planes cut Δ. It follows that if the depth is measured with respect to weights $\hat{x}_i = 2x_i$, the new depth is at least $2\mathsf{d}(\Delta) - 1/2 > \mathsf{d}(\Delta)$. That is, the weights $\hat{x}_i$ give a feasible fractional solution to the multi-cover instance $(\Gamma, \mathcal{F}')$. Note that since $\beta < 1/4$, the weights $\hat{x}_i$ satisfy $\hat{x}_i < 1$.

The remainder of the algorithm is to apply the approach implied by Lemma 4.3: we find a $c\hat{x}$-sample $\mathcal{R}$, with c to be determined; this induces a residual multi-cover problem $(\Gamma, \mathcal{F}') \setminus \mathcal{R}$, which we solve using the simple technique of Lemma 4.1. Letting U denote the resulting combined solution to $(\Gamma, \mathcal{F}')$, we return $U \cup X$ as a cover for the original multi-cover problem.

The analysis of this algorithm is the proof of the following result.

Theorem 4.5 *Let* $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ *be an instance of multi-cover formed by a set* P *of points in* $\mathbb{R}^3$, *and a set* $\mathcal{F}$ *of halfspaces. Then, one can compute, in randomized polynomial time, a subset of halfspaces of* $\mathcal{F}$ *that meets all the required demands, and is of expected size* $O(\mathsf{f})$, *where* f *is the value of an optimal fractional solution to* LP.

Proof: We described the algorithm above, except for the values of c and β.

By Lemma 4.4, the expected number of simplices in the cutting Γ of demand at most t is $O\!\left(Wt^2\right)$, where $W = \mathsf{f}' \leq \mathsf{f}(\mathcal{I})$, which implies that Lemma 4.3 can be applied, with $V = \mathsf{f}(\mathcal{I})$, K an absolute constant, and using the weights $\hat{x}_i$. Since $\sum_i \hat{x}_i \leq 2\mathsf{f}(\mathcal{I})$, the expected size of U is at most $(c+2)\mathsf{f}(\mathcal{I})$, using the absolute constant value of c used in this application of Lemma 4.3. Observing that $|X| \leq \mathsf{f}(\mathcal{I})/\beta$, and taking $\beta = 1/2c$ to allow the cx-sample probabilities $c\hat{x}_i$ to be less than one, we have that the returned solution $U \cup X$ to $\mathcal{I}$ has expected cardinality at most $(3c+2)\mathsf{f}(\mathcal{I})$, which is $O(\mathsf{f}(\mathcal{I}))$.

The only non-trivial step in terms of verifying the running time is for computing the cutting and Lemma 4.4 guarantees the running time. ∎

Remark 4.6 The shallow-cutting lemma (Lemma 4.4) is shown via a random sampling argument, and our rounding algorithm is also based on random sampling, given the cutting as a black-box. One could do a direct analysis of random sampling by unfolding the proof of the cutting lemma. However, the indirect approach is easier to see and highlights the intuition behind the proof.

5 Generalizations and Applications

We now examine to what extent the result derived for covering points in $\mathbb{R}^3$ by halfspaces generalizes to other shapes.

5.1 Well behaved shapes

We are interested in set systems $(\mathsf{P}, \mathcal{F})$ where $\mathcal{F}$ is a set of "well-behaved" shapes such as disks or fat triangles. As we remarked already, it is shown in [CV07] that the existence of good ε-nets for such shapes can be derived from bounds on their union complexity. For example, it is shown that if $\mathcal{F}$ is a set of *fat* triangles in the plane then there is an $O(\log \log \mathsf{f})$ approximation for the set cover problem. For fat wedges one obtains an $O(1)$ approximation. Here we show that union complexity bounds can be used to derive approximation ratios for the multi-cover problem that are similar to those derived in [CV07] for the set cover problem. Following the scheme for halfspaces, the key tool is the existence of shallow cuttings. To this end we describe some general conditions for the shapes of interest and then state a shallow cutting lemma.

Let $\mathcal{F}$ be a set of n shapes in $\mathbb{R}^d$, such that their union complexity for any subset of size r is (at most) $\mathcal{U}(r)$, for some function $\mathcal{U}(r) \geq r$. Similarly, let $O(r^d)$ be the upper bound on total complexity of an arrangement of r such shapes.

Let X be a subset of $\mathbb{R}^d$. We assume that given a subset $\mathcal{G} \subseteq \mathcal{F}$, one can perform a decomposition the faces of the arrangement $\mathcal{A}(\mathcal{G})$ that intersects X into cells of constant descriptive complexity (e.g., vertical trapezoids), and the complexity of this decomposition is proportional to the number of vertices of the faces of $\mathcal{A}(\mathcal{G})$ that intersects X. Finally, we assume that the intersection of d shapes of $\mathcal{F}$ generates a constant number of vertices.

One can then derive the following version of Matoušek's shallow cutting lemma. We emphasize that this lemma is a straightforward (if slightly messy) adaption of the result of Matoušek. A proof is sketched in Appendix A.

Theorem 5.1 *Given a set $\mathcal{F}$ of "well-behaved" shapes in $\mathbb{R}^d$ with total weight n, and parameters r and k, one can compute a decomposition of space into $O(r^d)$ cells of constant descriptive complexity, such that total weight of boundaries of shapes of $\mathcal{F}$ intersecting a single cell is at most n/r. Furthermore, the expected total number of cells containing points of depth smaller than k is*

$$
O\left(\left(\frac{rk}{n} + 1 \right)^d \mathcal{U}\left(\frac{n}{k} \right) \right),
$$

where $\mathcal{U}(\ell)$ is the worst-case combinatorial complexity of the boundary of the union of ℓ shapes of $\mathcal{F}$.

Using the same scheme as that for halfspaces we can derive approximation ratios for the multi-cover problem for shapes that have the property that $\mathcal{U}(n)$ is near-linear in n. An approximation ratio of $O(\mathcal{U}(\text{opt}) / \text{opt})$ easily follows, but in fact, by using the oversampling idea of Aronov et al. [AES09], we can improve this to $O(\log(\mathcal{U}(\text{opt}) / \text{opt}))$. We use the shallow cutting lemma as a black box, and hence our argument is arguably slightly simpler than then one in [AES09] and our result can be interpreted as a generalization.

Theorem 5.2 *Let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be an instance of multi-cover formed by a set P of points in $\mathbb{R}^d$, and a set $\mathcal{F}$ of ranges. Furthermore, the union complexity of any ℓ such ranges is (at most) $\mathcal{U}(\ell)$, for some function $\mathcal{U}(\ell) \geq \ell$. Then, one can compute, in randomized polynomial time, a subset of ranges of $\mathcal{F}$ that meets all the required demands, and is of expected size $O\left(\mathsf{f} \log \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}} \right)$, where f is the value of an optimal fractional solution to LP.*

Proof: As before, we compute the LP relaxation, and take all the ranges that the value of $x_i \geq \beta$, where $\beta = \alpha / \log \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}}$ for some sufficiently small constant α. Next, we compute a $(1/4\mathsf{f})$-cutting Γ of residual system $(\mathsf{P}', \mathcal{F}')$. Using Theorem 5.1 with parameters $r = 4\mathsf{f}$, $n = \mathsf{f}$ and $k = t$, there are at most

$$
O\left((t+1)^d \, \mathsf{f} \, \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}} \right)
$$

cells, with depth at most t. In particular, this bounds the number of cells in the cutting with depth in the range $t-1$ to t. We pick a random sample $\mathcal{R}$ of (expected) size $h = O\left(\mathsf{f} \log \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}} \right)$ from $\mathcal{F}'$, by performing a cx-sample from $\mathcal{F}'$, where $c = O\left(\log \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}} \right)$. Arguing as in Lemma 4.2, the expected residual demand for a cell of Γ with demand t is $t \exp(-ct/4)$. Therefore, the expected total residual

demand in $(\Gamma, \mathcal{F}') \setminus \mathcal{R}$ is

$$O\left(\sum_{t=1}^{\infty} \exp\left(-\frac{c}{4}t\right)(t+1)^{d+1}\, \mathsf{f}\, \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}}\right) = O(\mathsf{f}).$$

Using Lemma 4.1, the residual multi-cover instance $(\Gamma, \mathcal{F}') \setminus \mathcal{R}$ has a cover of expected size $O(\mathsf{f})$. Thus, we have shown that the original multi-cover instance has a cover of expected size $O(\mathsf{f}/\beta + h + \mathsf{f}) = O\left(\mathsf{f} \log \frac{\mathcal{U}(\mathsf{f})}{\mathsf{f}}\right).$ ∎

Applications: The above general result can be combined with known bounds on $\mathcal{U}(n)$ to give several new results. We follow [CV07, AES09] who gave approximation ratios for the set cover problem using a similar general framework; we give essentially similar bounds for the multi-cover problem. All the instances below involve shapes in the Euclidean plane.

- $O(1)$ approximation for pseudo-disks, fat triangles of similar size, and fat wedges.

- $O(\log \log \log \mathsf{f})$ approximation for fat triangles (which also implies similar bounds for fat convex polygonal shapes of constant description complexity).

- $O(\log \alpha(\mathsf{f}))$ approximation for regions each of which is defined by the intersection of the non-negative y halfplane with a Jordan region such that each pair of bounding Jordan curves intersecting at most three times (not counting the intersections on the x axis). Here $\alpha(n)$ is the inverse Ackerman function.

5.2 Unit Cubes in 3d

We also get a similar result for the case of axis-parallel unit cubes.

In [CV07] an $O(1)$ approximation is also shown that for the problem of covering points by unit sized axis parallel cubes in three dimensions. There is a technical difficulty for this case. Although it is known from [BSTY98] that the combinatorial complexity of the union of n cubes is $O(n)$, the same bound is not known for the canonical decomposition of the exterior of the union as required by our framework. The same difficulty is present in [CV07] and they overcome this by taking advantage of the fact that all cubes are unit sized. The basic idea is to use a grid shifting trick to decompose the given instance into independent instances such that each instances has cubes that contain a common intersection point. For this special case one can show that the canonical decomposition of the exterior of the shapes has linear complexity. This suffices for the framework in [CV07]. For our framework we need a cutting.

Lemma 5.3 *Let S be a set of n axis-parallel unit cubes in three dimensions, all of them containing (say) the origin. Then, one can decompose the arrangement of $\mathcal{A}(S)$ into a canonical decomposition of axis parallel boxes, such that the complexity of decomposing every face is proportional to the number of vertices on its boundary.*

Proof: First we break the arrangement into eight octants by the three axis planes (xy, yz and xz planes). We will describe how to decompose the arrangement in the positive octant, and by symmetry the construction would apply to the whole arrangement.

So, let f be a 3d face of the arrangement (when clipped to the positive octant). Let I be the cubes of S that contain f, and similarly, let B be the set of cubes of S that contribute to the

boundary of f, but do not include f in their interior. As such, we have that

$$f = \text{closure}\left(\left(\bigcap_{c \in I} c\right) \setminus \left(\bigcup_{c' \in B} c'\right)\right).$$

(If the set I is empty, we will add a fake huge cube to ensure f is bounded.) Now, the first term is just an axis-parallel box. Intuitively, the second term (the "floor" of f) is a (somewhat bizarre) collection of "stairs". Note, that any vertical line that intersects f, intersects it in an interval. In particular, let g the top face (in the z direction) of f, and observe that, since all the cubes of S contain the origin, it must be that any line that intersect f must also intersect g. As such, let us project all the edges and vertices of f upward till the hit g. This results in a collection W of (interior) disjoint segments that partition (the rectangular polygon) f. We perform a vertical decomposition of the 2d arrangement formed by $\mathcal{A}(W)$ (including the outer face of this arrangement, which is g). This results in $O(|f|)$ collection of (interior) disjoint rectangles that cover g, where $|f|$ is the number of vertices on the boundary of f. Furthermore, for such a rectangle r, there is no edge or vertex of f, such that their vertical projection lies in the interior of r. Namely, we can erect a vertical prism for each face of the vertical decomposition of $\mathcal{A}(W)$, till the prism hits the bottom boundary of f. This result in a decomposition of f into $O(|f|)$ disjoint boxes, as required. ∎

Lemma 5.3 implies that an the arrangement $\mathcal{A}(S)$, can be decomposed into (canonical) boxes, in such a way that the number of boxes of certain depth t, is proportional to the number of vertices of $\mathcal{A}(S)$ of this depth. This implies that we can apply the shallow cutting lemma to S (we remind the reader that all the axis-parallel unit cubes of S contain the origin).

This is sufficient to imply $O(1)$ approximation to multi-cover. Indeed, let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be the given instance of multi-cover, where $\mathcal{F}$ is a set of unit-cubes in 3d. Let G be the unit grid, and for any point $\mathsf{q} \in G$, let $\mathcal{F}_\mathsf{q}$ be the set of cubes of $\mathcal{F}$ that contains p (for the simplicity of exposition, we assume that every cube of $\mathcal{F}$ is contained in exactly one such set, as this can be easily guaranteed by shifting G slightly). Next, solve the LP associated with $\mathcal{I}$, and associate a point $\mathsf{p} \in \mathsf{P}$ with $\mathsf{q} \in G$, if the depth of p in $\mathcal{F}_\mathsf{q}$ is at least $1/8$ (if p can be associated with several such instances, we pick the one that provides maximum coverage for p). Let P_q be the resulting set of points. Thus, for any point in $\mathsf{q} \in G$, there is an associated instance of multi-cover $(\mathsf{P}_\mathsf{q}, \mathcal{F}_\mathsf{q})$. Clearly, a constant factor approximation for each of these instances, would lead to a constant factor approximation for the whole problem.

Now, $\mathcal{F}_\mathsf{q}$ is made of cubes all containing a common point, and as such Lemma 5.3 implies that shallow cutting would work for it. In particular, we can now apply the algorithm of Theorem 4.5 to this instance, and get a constant factor approximation (here, implicitly, we also used the fact that the union complexity of n axis-parallel unit cubes is linear). This implies the following theorem.

Theorem 5.4 *Let $\mathcal{I} = (\mathsf{P}, \mathcal{F})$ be an instance of multi-cover formed by a set P of points in $\mathbb{R}^3$, and a set $\mathcal{F}$ of axis-parallel unit cubes. Then, one can compute, in randomized polynomial time, a subset of cubes of $\mathcal{F}$ that meets all the required demands, and is of expected size $O(\mathsf{f})$, where f is the value of an optimal fractional solution to LP.*

6 Conclusions

We presented improved approximation algorithms for set multi-cover in geometric settings. Our key insight was to produce a "small" instance of the problem by clustering the given instance. This in turn was done by using a variant of shallow cuttings. We believe that this approach might be useful for other problems in geometric settings.

An interesting open problem, is to obtain improved algorithms for the set cover and the set multi-cover problems in geometric settings when the sets/shapes have costs associated with them and the goal is to find a cover of lowest cost. Can the results from [Cla93, BG95, CV07] and this paper be extended to this more general setting?

References

[AES09]   B. Aronov, E. Ezra, and M. Sharir. Small-size eps-nets for axis-parallel rectangles and boxes. In *Proc. 41st Annu. ACM Sympos. Theory Comput.*, 2009.

[BG95]   H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:263–279, 1995.

[BSTY98]   J. D. Boissonnat, M. Sharir, B. Tagansky, and M. Yvinec. Voronoi diagrams in higher dimensions under certain polyhedral distance functions. *Discrete Comput. Geom.*, 19(4):485–519, 1998.

[CF90]   B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.

[Cla88]   K. L. Clarkson. Applications of random sampling in computational geometry, II. In *Proc. 4th Annu. ACM Sympos. Comput. Geom.*, pages 1–11, 1988.

[Cla93]   K. L. Clarkson. Algorithms for polytope covering and approximation. In *Proc. 3th Workshop Algorithms Data Struct.*, volume 709 of *Lect. Notes in Comp. Sci.*, pages 246–252. Springer-Verlag, 1993.

[CS89]   K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.

[CV07]   K. L. Clarkson and K. R. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37(1):43–58, 2007.

[dBS95]   M. de Berg and O. Schwarzkopf. Cuttings and applications. *Internat. J. Comput. Geom. Appl.*, 5:343–355, 1995.

[ERS05]   G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Inform. Process. Lett.*, 95(2):358–362, 2005.

[Fei98]   Uriel Feige. A threshold of $\ln n$ for approximating set cover. *J. Assoc. Comput. Mach.*, 45(4):634–652, 1998.

[FG88]   T. Feder and D. H. Greene. Optimal algorithms for approximate clustering. In *Proc. 20th Annu. ACM Sympos. Theory Comput.*, pages 434–444, 1988.

[FMZ07]   C. Fragoudakis, E. Markou, and S. Zachos. Maximizing the guarded boundary of an art gallery is apx-complete. *Comput. Geom. Theory Appl.*, 38(3):170–180, 2007.

[Har08]   S. Har-Peled. Geometric approximation algorithms. Class notes. Online at `http://uiuc.edu/~sariel/teach/notes/aprx/`, 2008.

[LLS01]   Y. Li, P. M. Long, and A. Srinivasan. Improved bounds on the sample complexity of learning. *J. Comput. Syst. Sci.*, 62(3):516–527, 2001.

[Lon01] P. M. Long. Using the pseudo-dimension to analyze approximation algorithms for integer programming. In *Proc. 7th Workshop Algorithms Data Struct.*, volume 2125 of *Lecture Notes Comput. Sci.*, pages 26–37, 2001.

[LY94] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *J. Assoc. Comput. Mach.*, 41(5):960–981, 1994.

[Mat92] J. Matoušek. Reporting points in halfspaces. *Comput. Geom. Theory Appl.*, 2(3):169–186, 1992.

[MR95] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, 1995.

[PA95] J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.

[TWDJ08] M. T. Thai, F. Wang, D. H. Du, and X. Jia. Coverage problems in wireless sensor networks: design and analysis. *Inter. J. Sensor Networks*, 3(3):191–200, 2008. Special issue on Coverage Problems in Sensor Networks.

[Var09] K. Varadarajan. Epsilon nets and union complexity. In *Proc. 25th Annu. ACM Sympos. Comput. Geom.*, 2009.

[Vaz01] V. V. Vazirani. *Approximation algorithms*. Springer-Verlag New York, Inc., New York, NY, USA, 2001.

A   A shallow cutting lemma for "nice" shapes

In this section, we prove Theorem 5.1, a variant of the shallow cutting lemma of Matoušek in a slightly different setting. We include the details for the sake of completeness, which are not hard in light of Matoušek's work [Mat92]. Our description is somewhat informal, for simplicity. The family of shapes that we consider needs to satisfy the assumptions outlined in Section 5.

Building $(1/r)$-cuttings.   When computing cuttings, one first picks a random sample $\mathcal{R}$ of size r of the objects of $\mathcal{F}$, and computes the decomposition $\mathcal{A}_{\parallel}(\mathcal{R})$ of the arrangement of the random sample. For a cell $\triangle$ in this decomposition, let $\mathrm{cl}(\triangle)$ be the list of shapes of $\mathcal{F}$ whose boundaries intersect the interior of $\triangle$. If $|\mathrm{cl}(\triangle)| \leq n/r$ then it is acceptable, and we add it to the resulting cutting.

Otherwise, we need to do a local patching up, by partitioning each such cell further. Specifically, let $t_{\triangle} = \lceil \mathrm{cl}(\triangle)/(n/r) \rceil$ be the ***excess*** of $\triangle$. We take a random sample $\mathcal{R}_{\triangle}$ of size $O(t_{\triangle} \log(t_{\triangle}))$ from $\mathrm{cl}(\triangle)$. With constant probability, this is a $1/t_{\triangle}$-net of $\mathrm{cl}(\triangle)$ (for ranges formed by our decomposition). We verify that it is such a net, and if not, we resample, and repeatedly do so until we obtain a $1/t_{\triangle}$-net. To do the verification, we build the arrangement of $\mathcal{R}_{\triangle}$ inside $\triangle$, and compute its decomposition, and check that all the cells in this decomposition intersect at most n/r boundaries of the shapes of $\mathcal{F}$. Let $\mathrm{dcmp}(\triangle)$ denote this decomposition of $\triangle$ (if $\triangle$ has excess at most 1, then we just take $\mathrm{dcmp}(\triangle)$ to be $\{\triangle\}$). Clearly, the set

$$\bigcup_{\triangle \in \mathcal{A}_{\parallel}(\mathcal{R})} \mathrm{dcmp}(\triangle)$$

forms a decomposition of $\mathbb{R}^d$ into regions of constant complexity, and each region intersects at most n/r boundaries of the shapes of $\mathcal{F}$.

It is well known that the complexity of the resulting cutting is (in expectation) $O(r^d)$ [CF90]

Let $\mathcal{C}$ denote the resulting cutting.

Size of cutting at a certain depth. Here we are interested in the number of cells in the arrangement $\mathcal{A}_{\parallel}(\mathcal{R})$ that cover "shallow" portions of $\mathcal{A}(\mathcal{F})$. Formally, the ***depth*** of a point $\mathsf{p} \in \mathbb{R}^d$, is the number of shapes of $\mathcal{F}$ that cover it. Let $f_{\leq k}(n)$ denote the maximum number of vertices of depth at most k in an arrangement of n shapes. Clarkson and Shor [CS89] showed that $f_{\leq k}(n) = O\bigl(k^d \mathcal{U}(n/k)\bigr)$. Specifically, we are interested in the number of cells of $\mathcal{C}$ that contain points of depth at most k. The ***kth level*** is the closure of all the points on the boundary of the shapes that are contained inside k shapes.

Now, the expected number of vertices of $\mathcal{A}(\mathcal{R})$ that are of depth at most k in $\mathcal{A}(\mathcal{F})$ is

$$ O\left(\left(\frac{r}{n}\right)^d k^d \mathcal{U}(n/k)\right) = O\left(\left(\frac{rk}{n}\right)^d \mathcal{U}(n/k)\right), $$

since for a given vertex of $\mathcal{A}(\mathcal{F})$ of depth at most k, the probability that all d shapes that define it will picked to be in $\mathcal{R}$ is $O\bigl((r/n)^d\bigr)$. This unfortunately does not bound the number of cells in the decomposition of $\mathcal{A}_{\parallel}(\mathcal{R})$ that contain points of depth at most k, since we might have cells that cross the kth level.

So, let $X \subseteq \mathbb{R}^d$ be a fixed subset of space, and let $x(|\mathcal{R}|)$ be the number of cells of $\mathcal{A}_{\parallel}(\mathcal{R})$ that intersect X. Let $x(r)$ denote the maximum value of $x(|\mathcal{R}|)$ over all samples $\mathcal{R}$ of size r. Similarly, let $x^t(\mathcal{R})$ denote the number of cells in $\mathcal{A}_{\parallel}(\mathcal{R})$ that intersect X and have excess more than t (i.e., there are at least $t \cdot n/r$ shapes intersecting this cell).

Chazelle and Friedman [CF90] showed an exponential decay lemma stating that $\mathbf{E}\bigl[x^t(\mathcal{R})\bigr] = O\bigl(2^{-t}\mathbf{E}[x(\mathcal{R})]\bigr)$. We comment that, in fact, one can prove directly from the Clarkson-Show technique a polynomial decay lemma, which is sufficient to prove the shallow-cutting lemma. This polynomial decay lemma is implicit in the work of de Berg and Schwarzkopf [dBS95] although it was not stated explicitly (it also made it stealthy appearance in Clarkson and Varadarajan work [CV07], but [dBS95] seems to be the earliest reference).

Lemma A.1 (Polynomial decay lemma.) *For $t \geq 1$, let $\mathcal{R}$ be a random sample of size r from $\mathcal{F}$, and let $c \geq 1$ be an arbitrary constant. Then $\mathbf{E}\bigl[x^t(\mathcal{R})\bigr] = O(x(r)/t^c)$.*

Proof: By the Clarkson-Shor technique [CS89, Cla88], we have that

$$ \mathbf{E}\left[\sum_{\triangle \in \mathcal{A}_{\parallel}(\mathcal{R})} |\mathrm{cl}(\triangle)|^c\right] = O\left(\left(\frac{n}{r}\right)^c \mathbf{E}[x(\mathcal{R})]\right). = O\left(\left(\frac{n}{r}\right)^c x(r)\right). $$

In particular, if there are $x^t(\mathcal{R})$ cells in $\mathcal{A}_{\parallel}(\mathcal{R})$ with conflict-list of size larger than $t(n/r)$, then they contribute to the left size of the above equation the quantity $x^t(\mathcal{R})(t(n/r))^c$. We conclude that

$$ \mathbf{E}\bigl[x^t(\mathcal{R})(t(n/r))^c\bigr] = O\left(\left(\frac{n}{r}\right)^c x(r)\right), $$

which implies that $\mathbf{E}\bigl[x^t(\mathcal{R})\bigr] = O(x(r)/t^c)$, as claimed. ∎

Lemma A.2 *The expected number of cells in the $(1/r)$-cutting $\mathcal{C}$ of $\mathcal{F}$ that contain points of depth at most k is bounded by*

$$O\left(\left(\frac{rk}{n}+1\right)^d \mathcal{U}\left(\frac{n}{k}\right)\right).$$

Proof: If a cell $\triangle$ of $\mathcal{A}_{\|}(\mathcal{R})$ has excess t, and it intersects the kth level, then all its points have depth at most $k + t(n/r)$. The expected number of vertices of $\mathcal{A}_{\|}(\mathcal{R})$ of depth at most $\alpha(t) = k + t(n/r)$ is

$$\gamma(t) = O\left(\left(\frac{r\alpha(t)}{n}\right)^d \mathcal{U}\left(\frac{n}{\alpha(t)}\right)\right)$$

which also (asymptotically) bounds the number of cells in $\mathcal{A}_{\|}(\mathcal{R})$ having depth smaller than $\alpha(t)$. Let X_t denote the number of cells with excess t (or more) with depth at most $\alpha(t)$. Setting $c = O(d)$, we have by the polynomial decay lemma, that

$$\mathbf{E}[X_t] = O\left(\gamma(t)/t^{4d}\right) = O\left(\left(\frac{r\alpha(t)}{t^4 n}\right)^d \mathcal{U}\left(\frac{n}{\alpha(t)}\right)\right).$$

Now, the number of cells of the cutting $\mathcal{C}$ that have points with depth at most k is bounded by

$$Y = O\left(\sum_{t=0}^{\infty} X_t \cdot (t \log t)^d\right)$$

Thus, we have

$$
\begin{aligned}
\mathbf{E}[Y] &= O\left(\sum_{t=0}^{\infty} \mathbf{E}[X_t] \cdot t^{O(d)}\right) = O\left(\sum_{t=0}^{\infty} \left(\frac{r\alpha(t)}{t^c n}\right)^d \mathcal{U}\left(\frac{n}{\alpha(t)}\right) \cdot t^{O(d)}\right) \\
&= O\left(\left(\frac{r}{n}\right)^d \mathcal{U}\left(\frac{n}{k}\right) \sum_{t=0}^{\infty} t^{O(d)-c}\left(k + t(n/r)\right)^d\right) = O\left(\left(\frac{r}{n}\right)^d \mathcal{U}\left(\frac{n}{k}\right)(k + n/r)^d \sum_{t=0}^{\infty} t^{O(d)-c}\right) \\
&= O\left(\left(\frac{kr}{n}+1\right)^d \mathcal{U}\left(\frac{n}{k}\right)\right),
\end{aligned}
$$

by setting c to be sufficiently large. ∎

This proves Theorem 5.1 by using replication to represent weights.