Finding Small Hitting Sets in Infinite Range Spaces of Bounded VC-Dimension

Khaled Elbassioni

Masdar Institute of Science and Technology, Abu Dhabi, UAE
kelbassioni@masdar.ac.ae

Abstract

We consider the problem of finding a small hitting set in an *infinite* range space $\mathcal{F} = (Q, \mathcal{R})$ of bounded VC-dimension. We show that, under reasonably general assumptions, the infinite-dimensional convex relaxation can be solved (approximately) efficiently by multiplicative weight updates. As a consequence, we get an algorithm that finds, for any $\delta > 0$, a set of size $O(s_{\mathcal{F}}(z^*_{\mathcal{F}}))$ that hits $(1-\delta)$-fraction of $\mathcal{R}$ (with respect to a given measure) in time proportional to $\log(\frac{1}{\delta})$, where $s_{\mathcal{F}}(\frac{1}{\epsilon})$ is the size of the smallest ϵ-net the range space admits, and $z^*_{\mathcal{F}}$ is the value of the *fractional* optimal solution. This *exponentially* improves upon previous results which achieve the same approximation guarantees with running time proportional to $\text{poly}(\frac{1}{\delta})$. Our assumptions hold, for instance, in the case when the range space represents the *visibility* regions of a polygon in $\mathbb{R}^2$, giving thus a deterministic polynomial-time $O(\log z^*_{\mathcal{F}})$-approximation algorithm for guarding $(1-\delta)$-fraction of the area of any given simple polygon, with running time proportional to $\text{polylog}(\frac{1}{\delta})$.

1998 ACM Subject Classification G.1.6 Convex Programming, I.3.5 Geometric Algorithms, Languages, and Systems

Keywords and phrases VC-dimension, approximation algorithms, fractional covering, multiplicative weights update, art gallery problem, polyhedral separators, geometric covering

Digital Object Identifier 10.4230/LIPIcs.SoCG.2017.40

1 Introduction

Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space defined by a (possibly) *infinite* set of ranges $\mathcal{R} \subseteq 2^Q$ over a (possibly) *infinite* set Q. A *hitting* set of $\mathcal{R}$ is a subset $H \subseteq Q$ such that $H \cap R \neq \emptyset$ for all $R \in \mathcal{R}$. Finding a hitting set of minimum size for a given range space is a fundamental problem in computational geometry. For finite range spaces (that is, when Q is finite), standard algorithms for SetCover [29, 33, 13] yield $(\log |Q|+1)$-approximation in polynomial time, and this is essentially the best possible guarantee assuming $P \neq NP$ [17]. Better approximation algorithms exist for special cases, such as range spaces of *bounded VC-dimension* [8], of *bounded union complexity* [15, 47], of *bounded shallow cell complexity* [9], as well as several classes of geometric range spaces [3, 40, 30]. Many of these results are based on showing the existence of a small-size *ϵ-net* for the range space $\mathcal{F}$ and then using the multiplicative weights update algorithm of Brönnimann and Goodrich [8]. For instance, if a range space $\mathcal{F}$ has VC-dimension d, then it admits an ϵ-net of size $O(\frac{d}{\epsilon}\log\frac{1}{\epsilon})$ [27, 31], which by the above mentioned method implies an $O(d \cdot \log \text{Opt}_{\mathcal{F}})$-approximation algorithm for the hitting set problem for $\mathcal{F}$, where $\text{Opt}_{\mathcal{F}}$ denotes the size of a minimum-size hitting set. Even et al. [19] observed that this can be improved to $O(d \cdot \log z^*_{\mathcal{F}})$-approximation by first solving the LP-relaxation of the problem to obtain the value of the *fractional* optimal solution $z^*_{\mathcal{F}}$, and then finding an ϵ-net, with $\epsilon := 1/z^*_{\mathcal{F}}$.

33rd International Symposium on Computational Geometry (SoCG 2017).
Editors: Boris Aronov and Matthew J. Katz; Article No. 40; pp. 40:1–40:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The multiplicative weights update algorithm in [8] works by maintaining weights on the *points*. The straightforward extension to infinite (or continuous) range spaces (that is, the case when Q is infinite) does not seem to work, since the bound on the number of iterations depends on the measure of the regions created during the course of the algorithm, which can be arbitrarily small. In this paper we take a different approach, which can be thought of as a combination of the methods in [8] and [19] (with LP replaced by an *infinite-dimensional convex relaxation*[1]):

- We maintain weights on the *ranges* (in contrast to the method of Brönnimann and Goodrich [8] which maintains weights on the points, and the second method suggested by Agarwal and Pan [1] which maintains weights on both points and ranges);
- We first solve the covering convex relaxation within a factor of $1+\varepsilon$ using the multiplicative weights update (MWU) methoid, extending the approach in [21] to infinite-dimensional covering LP's (under reasonable assumptions);
- We finally use the rounding idea of [19] to get a small integral hitting set from the obtained fractional solution.

Informal main theorem. There is an algorithm that, given a range space $\mathcal{F} = (Q, \mathcal{R})$ of VC-dimension d and $\delta > 0$, (under mild assumptions) finds a subset of Q of size $O(d \cdot z^*_{\mathcal{F}} \log z^*_{\mathcal{F}})$ that hits $(1-\delta)$-fraction of $\mathcal{R}$ (with respect to a given measure) in time polynomial in the input description of $\mathcal{F}$ and $\log(\frac{1}{\delta})$.

This *exponentially* improves upon previous results[2] which achieve the same approximation guarantees, but with running time depending *polynomially* on $\frac{1}{\delta}$. It should be noted that the main contribution of this paper is an efficient algorithm for approximately solving the *infinite-dimensional* covering linear programming relaxation with an *infinite* number of constraints. Even though such relaxation is convex, none of the known polynomial-time methods for convex programming (such as the ellipsoid method and interior point methods) can be used since their running time depends polynomially on the dimension. For the class of infinite-dimensional covering linear programs with infinitely many constraints corresponding to the ranges of a range space, we observe an interesting connection between the convergence of the MWU method and the fact that the range space has bounded VC-dimension.

We apply this result to a number of problems:

- The art gallery problem: given a simple polygon H, our main theorem implies that there is a *deterministic* polytime $O(\log z^*_{\mathcal{F}})$-approximation algorithm (with running time proportional to $\mathrm{polylog}(\frac{1}{\delta})$) for guarding $(1-\delta)$-fraction of the area of H. When δ is (exponentially) small, this improves upon a previous result [12] which gives a polytime algorithm that finds a set of size $O(\textsc{Opt}_{\mathcal{F}} \cdot \log \frac{1}{\delta})$, guarding $(1-\delta)$-fraction of the area of H. Other (*randomized*) $O(\log \textsc{Opt}_{\mathcal{F}})$-approximation results which provide full guarding (i.e., $\delta = 0$) also exist, but they either run in pseudo-polynomial time [16], restrict the set of candidate guard locations [18, 22], or make some general position assumptions [6].
- Covering a polygonal region by translates of a convex polygon: Given a collection $\mathcal{H}$ of polygons in the plane and a convex polygon H_0, our main theorem implies that there is a randomized polytime $O(1)$-approximation algorithm for covering $(1-\delta)$ of the total area

[1] More precisely, an *infinite-dimensional* covering linear programming relaxation with an *infinite* number of constraints

[2] More precisely (as pointed to us by an anonymous reviewer), using relative approximation results (see, e.g., [26]), one can obtain the same approximation guarantees as our main Theorem by solving the problem on the set system induced on samples of size $O((d \cdot \textsc{Opt}_{\mathcal{F}}/\delta) \log(1/\delta))$

of the polygons in $\mathcal{H}$ by the minimum number of translates of H_0. Previous results with proved approximation guarantees mostly consider only the case when $\mathcal{H}$ is a set of points [15, 28, 32].

- Polyhedral separation in fixed dimension: Given two convex polytopes $\mathcal{P}_1, \mathcal{P}_2 \subseteq \mathbb{R}^d$ such that $\mathcal{P}_1 \subset \mathcal{P}_2$, our main theorem implies that there is a randomized polytime $O(d \cdot \log z^*_{\mathcal{F}})$-approximation algorithm for finding a polytope $\mathcal{P}_3$ with the minimum number of facets separating $\mathcal{P}_1$ from $(1-\delta)$-fraction of the volume of $\partial\mathcal{P}_2$. This improves the approximation ratio by a factor of d over the previous (deterministic) result [8] (but which gives a complete separation).

The paper is organized as follows. In the next section we define our notation, recall some preliminaries, and describe the infinite-dimensional convex relaxation. In Section 3, we state our main result, followed by the algorithm for solving the fractional problem in Section 4 and its analysis in Section 5. The success of the whole algorithm relies crucially on being able to efficiently implement the so-called *maximization oracle*, which essentially calls for finding, for a given measure on the ranges, a point that is contained in the heaviest subset of ranges (with respect to the given measure). We utilize the fact that the dual range space has bounded VC-dimension in Section 6 to give an efficient randomized implementation of the maximization oracle in the *Real RAM* model of computation. With more work, we show in fact that, in the case of the art gallery problem, the maximization oracle can be implemented in *deterministic* polynomial time in the *bit model*. Sections 7.2 and 7.3 describe the two other applications.

2 Preliminaries

2.1 Notation

Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space. For a point $q \in Q$ and a subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$, let $\mathcal{R}'[q] := \{R \in \mathcal{R}' : \ q \in R\}$. The *dual* range space $\mathcal{F}^* = (Q^*, \mathcal{R}^*)$ is defined as the range space with $Q^* := \mathcal{R}$ and $\mathcal{R}^* := \{\mathcal{R}[q] : \ q \in Q\}$. For a set of points $P \subseteq Q$, let $\mathcal{R}_{|P} := \{R \cap P \ : \ R \in \mathcal{R}\}$ be the *projection* of $\mathcal{R}$ onto P. Similarly, for a set of ranges $\mathcal{R}' \subseteq \mathcal{R}$, let $\mathcal{R}^*_{|\mathcal{R}'} := \{\mathcal{R}'[q] : \ q \in Q\}$. For a positive integer r, we denote by $g_{\mathcal{F}}(r) \le 2^r$ the smallest integer such that for every finite set $P \subseteq Q$ of size r, we have $|\mathcal{R}_{|P}| \le g_{\mathcal{F}}(r)$. For $p \in Q$ and $R \in \mathcal{R}$, we denote by $\mathbb{1}_{p\in R} \in \{0,1\}$ the indicator variable that takes value 1 if and only if $p \in R$.

2.2 Assumptions

We shall make the following assumptions:

(A1) $g_{\mathcal{F}}(r) \le cr^{\gamma}$, for all integers $r \ge 0$ and some constants $\gamma \ge 1$ and $c > 0$ (known to the algorithm).

(A2) There exists a finite integral optimum whose value $\textsc{Opt}_{\mathcal{F}}$ is bounded by a parameter n (that is not necessarily part of the input description).

(A3) There exists an integrable function $w_0 : \mathcal{R} \to \mathbb{R}_+$. We assume that the integration of w_0 over any subset of $\mathcal{R}$ of input description of size k can be computed in time poly(k).

Note that if assumption (A3) holds then w_0 naturally defines a finite measure on $\mathcal{R}$ where the measure for a (measurable) set $\mathcal{R}' \subseteq \mathcal{R}$ is $w_0(\mathcal{R}') := \int_{R\in\mathcal{R}'} w_0(R)dR$. [3]

[3] One may also consider a general measure on $\mathcal{R}$. For simplicity of presentation, we assume here the more restrictive condition (A3), as all the applications we consider have this restriction. However, the

2.3 Range spaces of bounded VC-dimension

We consider range spaces $\mathcal{F} = (Q, \mathcal{R})$ of bounded *VC-dimension* defined as follows. A finite set $P \subseteq Q$ is said to be *shattered* by $\mathcal{F}$ if $\mathcal{R}_{|P} = 2^P$. The VC-dimension of $\mathcal{F}$, denoted VC-dim$(\mathcal{F})$, is the cardinality of the largest subset of Q shattered by $\mathcal{F}$. If arbitrarily large subsets of Q can be shattered then VC-dim$(\mathcal{F}) = +\infty$. It is well-known [42, 43] that if VC-dim$(\mathcal{F}) = d$ then $g_{\mathcal{F}}(r) \le g(r, d) := \sum_{i=0}^{d} \binom{r}{i} = O(r^d)$, and that VC-dim$(\mathcal{F}^*) < 2^{d+1}$. Thus, if VC-dim$(\mathcal{F}) = d$ then Assumption (A1) is satisfied with $\gamma = d$.

2.4 ϵ-nets

Given a range space $(Q, \mathcal{R})$, an integrable function $\mu : Q \to \mathbb{R}_+$, and a parameter $\epsilon > 0$, an *ϵ-net* for $\mathcal{R}$ (w.r.t. μ) is a set $P \subseteq Q$ such that $P \cap R \neq \emptyset$ for all $R \in \mathcal{R}$ that satisfy $\mu(R) \ge \epsilon \cdot \mu(Q)$, where for $Q' \subseteq Q$, we write $\mu(Q') := \int_{q \in Q'} \mu(q)dq$. We say that a range space $\mathcal{F}$ admits an ϵ-net of size $s_{\mathcal{F}}(\cdot)$, if for any $\epsilon > 0$, there is an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$. For range spaces of VC-dimension d, it is known [27, 31] that a random sample (w.r.t. to the probability density function $\frac{\mu}{\mu(Q)}$) of size $s_{\mathcal{F}}(\frac{1}{\epsilon}) = O(\frac{d}{\epsilon} \log \frac{1}{\epsilon})$ is an ϵ-net with (high) probability $\Omega(1)$.

We say that $\mu : Q \to \mathbb{R}_+$ has (finite) support of size K if μ can be written as a conic combination of K *Dirac delta functions*[4]: for any $q \in Q$, $\mu(q) := \sum_{p \in P} \mu(p) \boldsymbol{\delta}_p(q)$, for some finite $P \subseteq Q$ of cardinality K and non-negative multipliers $\mu(p)$, for $p \in P$. If this is the case, an ϵ-net for $\mathcal{R}$ of size $s_{\mathcal{F}}(\frac{1}{\epsilon}) = O(\frac{d}{\epsilon} \log \frac{d}{\epsilon})$ can be computed deterministically in time $O(d)^{3d} \frac{1}{\epsilon^{2d}} \log^d(\frac{d}{\epsilon})K$ under the following assumption [7, 11, 34]:

(A1′) The range space is given by a *subsystem oracle* $\textsc{Subsys}(\mathcal{F}, P)$ that, given any finite $P \subseteq Q$, returns the set of ranges $\mathcal{R}_{|P}$ in time $O(|P|)^{d+1}$.

It should also be noted that some special range spaces may admit a smaller size ϵ-net, e.g., $s_{\mathcal{F}}(\frac{1}{\epsilon}) = O(\frac{1}{\epsilon})$ for half-spaces in $\mathbb{R}^3$ [35, 36]; see also [9, 30, 15, 47].

2.5 ϵ-approximations

Given the dual range space $\mathcal{F}^*$, a probability density function w on $\mathcal{R}$, and an $\epsilon > 0$, an ϵ-approximation is a finite subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$ such that, for all $q \in Q$,

$$\left| \frac{|\mathcal{R}'[q]|}{|\mathcal{R}'|} - \frac{w(\mathcal{R}[q])}{w(\mathcal{R})} \right| \le \epsilon; \tag{1}$$

see, e.g., [10]. It is known [2, 10, 46] that if $\mathcal{F} = (Q, \mathcal{R})$ is a range space of VC-dimension d, and w is an arbitrary probability density function on $\mathcal{R}$, then for any $\epsilon > 0$, a random sample (w.r.t. w) of size $O(\frac{d2^d}{\epsilon^2} \log \frac{1}{\epsilon\sigma})$ is an ϵ-approximation for $\mathcal{F}^*$, with probability $1 - \sigma$.

2.6 The fractional problem

Given a range space $\mathcal{F} = (Q, \mathcal{R})$, satisfying assumptions (A1)–(A3), the fractional covering problem for $\mathcal{F}$ seeks to find an integrable function $\mu : Q \to \mathbb{R}_+$, such that $\mu(R) \ge 1$ for all

extension to general measurable sets should be straightforward.

[4] A Dirac delta function satisfies $\int_{Q'} \boldsymbol{\delta}_p(q)dq = 1$ if $p \in Q'$, and $\int_{Q'} \boldsymbol{\delta}_p(q)dq = 0$, for any $Q' \subseteq Q$.

$R \in \mathcal{R}$ and $\mu(Q)$ is minimized[5]:

$$z^*_{\mathcal{F}} := \inf_{\text{integrable } \mu} \int_{q\in Q} \mu(q)dq \tag{F-HITTING}$$

$$\text{s.t.} \int_{q\in R} \mu(q)dq \geq 1, \qquad \forall R \in \mathcal{R}, \tag{2}$$

$$\mu(q) \geq 0, \qquad \forall q \in Q.$$

Equivalently, it is required to find a *probability* measure μ on Q that solves the maximin problem: $\sup_\mu \inf_{R\in\mathcal{R}} \mu(R)$.

▶ **Proposition 1.** *For a range space $\mathcal{F}$ satisfying (A2), we have $\text{OPT}_{\mathcal{F}} \geq z^*_{\mathcal{F}}$.*

Proof. Given a finite integral optimal solution P^*, we define an integrable function $\mu : Q \to \mathbb{R}_+$ of support size $\text{OPT}_{\mathcal{F}}$ by $\mu(q) := \sum_{p\in P^*} \boldsymbol{\delta}_p(q)$, for $q \in Q$. Then $\mu(Q) = \int_{q\in Q} \sum_{p\in P^*} \boldsymbol{\delta}_p(q)dq = \sum_{p\in P^*} \int_{q\in Q} \boldsymbol{\delta}_p(q)dq = \sum_{p\in P^*} 1 = |P^*| = \text{OPT}_{\mathcal{F}}$ and $\mu(R) = \int_{q\in R} \sum_{p\in P^*} \boldsymbol{\delta}_p(q)dq = \sum_{p\in P^*} \int_{q\in R} \boldsymbol{\delta}_p(q)dq = \sum_{p\in P^*} \mathbb{1}_{p\in R} = |\{p \in P^* : p \in R\}| \geq 1$, for all $R \in \mathcal{R}$, since P^* is a hitting set. Since μ is feasible for (F-HITTING), the claim follows. ◀

Assume $\mathcal{F}$ satisfies (A3). For $\alpha \geq 1$, we say that $\mu : Q \to \mathbb{R}_+$ is an α-*approximate* solution for (F-HITTING) if μ is feasible for (F-HITTING) and $\mu(Q) \leq \alpha \cdot z^*_{\mathcal{F}}$. For $\beta \in [0,1]$, we say that μ is β-feasible if $\mu(R) \geq 1$ for all $R \in \mathcal{R}'$, where $\mathcal{R}' \subseteq \mathcal{R}$ satisfies $w_0(\mathcal{R}') \geq \beta \cdot w_0(\mathcal{R})$. Finally, we say that μ is an (α, β)-approximate solution for (F-HITTING) if μ is both α-approximate and β-feasible.

2.7 Rounding the fractional solution

Brönnimann and Goodrich [8] gave a multiplicative weights update algorithm for approximating the minimum hitting set for a *finite* range space satisfying (A1′) and admitting an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$. Their algorithm works as follows. It first guesses the value of the optimal solution (within a factor of 2), and initializes the weights of all *points* to 1. It then finds an $\epsilon = \frac{1}{2\text{OPT}_{\mathcal{F}}}$-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$. If there is a range R that is not hit by the net (which can be checked by the subsystem oracle), the weights of all the points in R are doubled. The process is shown to terminate in $O(\text{OPT}_{\mathcal{F}} \log \frac{|Q|}{\text{OPT}_{\mathcal{F}}})$ iterations, giving an $s_{\mathcal{F}}(2\text{OPT}_{\mathcal{F}})/\text{OPT}_{\mathcal{F}}$-approximation. Even et al. [19] strengthen this result by using the linear programming relaxation to get $s_{\mathcal{F}}(z^*_{\mathcal{F}})/z^*_{\mathcal{F}}$-approximation. We can restate this result as follows.

▶ **Lemma 2.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space admitting an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$ and μ be a measure on Q satisfying (2). Then there is a hitting set for $\mathcal{R}$ of size $s_{\mathcal{F}}(\mu(Q))$.*

Proof. Let $\epsilon := \frac{1}{\mu(Q)}$. Then for all $R \in \mathcal{R}$ we have $\mu(R) \geq 1 = \epsilon \cdot \mu(Q)$, and hence an ϵ-net for $\mathcal{R}$ is actually a hitting set. ◀

▶ **Corollary 3.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d and $\mu : Q \to \mathbb{R}_+$ be an integrable function satisfying (2). Then a random sample of size $O(d \cdot \mu(Q) \log(\mu(Q)))$, w.r.t. the probability density function $\mu' := \frac{\mu}{\mu(Q)}$, is a hitting set for $\mathcal{R}$ with probability $\Omega(1)$. Furthermore, if μ has support size K then there is a deterministic algorithm that computes a hitting set for $\mathcal{R}$ of size $O(d \cdot \mu(Q) \log(d \cdot \mu(Q)))$ in time $O(d)^{3d} \mu(Q)^{2d} \log^d(d \cdot \mu(Q))K$.*

Further improvements on the Brönnimann-Goodrich algorithm can be found in [1].

[5] We may as well restrict μ to have finite support and replace the integrals over Q by summations.

3 Solving the fractional problem – Main result

We shall make the following further assumption:

(A4) There is a deterministic (resp., randomized) oracle $\textsc{Max}(\mathcal{F}, w, \nu)$ (resp., $\textsc{Max}(\mathcal{F}, w, \sigma, \nu)$), that given a range space $\mathcal{F} = (Q, \mathcal{R})$, an integrable function $w : \mathcal{R} \to \mathbb{R}_+$, and $\nu > 0$, returns (resp., with probability $1-\sigma$) a point $p \in Q$ such that $\xi_w(p) \geq (1-\nu)\max_{q\in Q}\xi_w(q)$, where $\xi_w(p) := w(\mathcal{R}[p]) = \int_{R\in\mathcal{R}} w(R)\mathbb{1}_{p\in R}dR$.

The following is the main result of the paper.

▶ **Theorem 4.** *Given a range space $\mathcal{F}$ satisfying (A1)–(A4), and $\varepsilon, \delta, \nu \in (0,1)$, there is a deterministic (resp., randomized) algorithm that finds (resp., with probability $\Omega(1)$) a function $\mu : Q \to \mathbb{R}_+$ of support size $K := O(\frac{\gamma}{\varepsilon^3(1-\nu)}\log\frac{\gamma}{\varepsilon}\cdot\textsc{Opt}_{\mathcal{F}}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\nu)})$ that is a $(\frac{1+5\varepsilon}{1-\nu}, 1-\delta)$-approximate solution for (F-hitting), using K calls to the oracle $\textsc{Max}(\mathcal{F}, w, \nu)$ (resp., $\textsc{Max}(\mathcal{F}, w, \sigma, \nu)$).*

In view of Corollary 3, we get the following theorem as an immediate consequence of Theorem 4.

▶ **Theorem 5** (Main Theorem). *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space satisfying (A1)–(A4) and admitting a hitting set of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$, and $\varepsilon, \delta, \nu \in (0,1)$ be given parameters. Then there is a (deterministic) algorithm that computes a set of size $s_{\mathcal{F}}(z^*_{\mathcal{F}})$, hitting a subset of $\mathcal{R}$ of measure at least $(1-\delta)w_0(\mathcal{R})$, using $O(\frac{\gamma}{\varepsilon^3(1-\nu)}\log\frac{\gamma}{\varepsilon}\cdot\textsc{Opt}_{\mathcal{F}}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\nu)})$ calls to the oracle $\textsc{Max}(\ldots, \nu)$ and a single call to an ϵ-net finder.*

In Section 6, we observe that the maximization oracle can be implemented in randomized polynomial time. As a consequence, we obtain the following corollary of Theorem 5, under the assumption of the availability of the following oracles:

- $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$: this is the dual subsystem oracle; given a finite subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$, it returns the set of ranges $\mathcal{R}^*_{|\mathcal{R}'}$. Note that $|\mathcal{R}^*_{|\mathcal{R}'}| \leq g(|\mathcal{R}'|, 2^{d+1})$.
- $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$: Given $\mathcal{F}$ and a finite subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$, the oracle returns a point $p \in Q$ that lies in $\bigcap_{R\in\mathcal{R}'} R$ (if one exists).
- $\textsc{Sample}(\mathcal{F}, \widehat{w})$: Given $\mathcal{F} = (Q, \mathcal{R})$ and a probability density function $\hat{w} : \mathcal{R} \to \mathbb{R}_+$, it samples a range $R \in \mathcal{R}$ according to $\widehat{w}$.

▶ **Corollary 6.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d satisfying (A2) and (A3) and $\varepsilon, \delta \in (0,1)$ be given parameters. Then there is a randomized algorithm that computes a set of size $O(d\cdot z^*_{\mathcal{F}}\log z^*_{\mathcal{F}})$, hitting a subset of $\mathcal{R}$ of measure at least $(1-\delta)w_0(\mathcal{R})$, in time $O(K\cdot(\tau_1\cdot N + \tau_2(N) + \tau_3(N) + g_{\mathcal{F}^*}(N)))$, where $K := O(\frac{d}{\varepsilon^3}\log\frac{d}{\varepsilon}\cdot\textsc{Opt}_{\mathcal{F}}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta})$, and τ_1, $\tau_2(N)$ and $\tau_3(N)$, are respectively the maximum times taken by the oracle $\textsc{Sample}(\mathcal{F}, \widehat{w})$, and the oracles $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$, $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ on a set $\mathcal{R}'$ of size $N := \frac{d2^d\textsc{Opt}^2_{\mathcal{F}}}{\varepsilon^2}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon}$.*

Note that $g_{\mathcal{F}^*}(r) \leq r^{2^{d+1}}$, but stronger bounds can be obtained for special cases.

4 The algorithm

The algorithm is shown in Algorithm 1 below. For any iteration t, let us define the *active* range-subspace $\mathcal{F}_t = (Q, \mathcal{R}_t)$ of $\mathcal{F}$, where

$$\mathcal{R}_t := \{R \in \mathcal{R} : |P_t \cap R| < T\}.$$

Data: A range space $\mathcal{F} = (Q, \mathcal{R})$ satsfying (A1)–(A4), and approximation accuracies $\varepsilon, \delta, \nu \in (0, 1)$.

Result: A $(\frac{1+5\varepsilon}{1-\nu}, 1-\delta)$-approximate solution μ for (F-HITTING).

1 $t \leftarrow 0$; $P_0 \leftarrow \emptyset$; set T as in (3)

2 **while** $w_0(\mathcal{R}_t) \geq \delta \cdot w_0(\mathcal{R})$ **do**

  `// define`[7] $w_t : \mathcal{R}_t \to \mathbb{R}_+$ `by` $w_t(R) \leftarrow (1-\varepsilon)^{|P_t \cap R|} w_0(R)$`, for` $R \in \mathcal{R}_t$

3   $p_{t+1} \leftarrow \text{MAX}(\mathcal{F}_t, w_t, \nu)$

4   $P_{t+1} \leftarrow P_t \cup \{p_{t+1}\}$

5   $t \leftarrow t + 1$

6 **end**

7 **return** the measure $\widehat{\mu} : Q \to \mathbb{R}_+$ defined by $\widehat{\mu}(q) \leftarrow \frac{1}{T} \sum_{p \in P_t} \boldsymbol{\delta}_p(q)$

Algorithm 1: The fractional covering algorithm

Clearly, (since these properties are hereditary) VC-dim$(\mathcal{F}_t) \leq$ VC-dim$(\mathcal{F})$, and $\mathcal{F}_t$ admits an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$ whenever $\mathcal{F}$ does. For convenience, we assume below that P_t is (possibly) a multi-set (repetitions allowed).

Define[6] $a := \frac{\gamma}{\varepsilon^2}$, $b := \max\{\ln T_0, 1\}$, $T_0 := \frac{\text{OPT}_{\mathcal{F}} \delta^{1/\gamma} c^{1/\gamma}}{\varepsilon(1-\nu)\delta^{1/\gamma}} \left(\ln \frac{1}{1-\varepsilon} + \ln \frac{1}{\varepsilon\delta} \right)$, and

$$T := e^2 ab(\ln(a + e - 1) + 1) = \Theta\left(\frac{\gamma}{\varepsilon^2} \log \frac{\gamma}{\varepsilon} \log \frac{\text{OPT}_{\mathcal{F}}}{\varepsilon\delta(1-\nu)} \right). \tag{3}$$

For simplicity of presentation, we will assume in what follows that the maximization oracle is deterministic; the extension to the probabilistic case is straightforward.

5 Analysis

Define the potential function $\Phi(t) := w_t(\mathcal{R}_t)$, where $w_t(R) := (1-\varepsilon)^{|P_t \cap R|} w_0(R)$, and $P_t = \{p_{t'} : t' = 1, \ldots, t\}$ is the set of points selected by the algorithm in Step 3 up to time t. We can also write $w_{t+1}(R) = w_t(R)(1 - \varepsilon \cdot \mathbb{1}_{p_{t+1} \in R})$.

The analysis is done in three steps: the first one (Section 5.1), which is typical for MWU methods, is to bound the potential function, at each iteration, in terms of the ratio between the current solution obtained by the algorithm at that iteration and the optimum fractional solution. The second step (Section 5.2) is to bound the number of iterations until the desired fraction of the ranges is hit. Finally, the third step (Section 5.3) uses the previous two steps to show that the algorithm reaches the required accuracy after a polynomial number of iterations.

5.1 Bounding the potential

The following three lemmas are obtained by the standard analysis of MWU methods with "$\sum$"'s replaced by "$\int$"'s.

▶ **Lemma 7.** *For all $t = 0, 1, \ldots,$ it holds that* $\Phi(t+1) \leq \Phi(t) \exp\left(-\frac{\varepsilon}{\Phi(t)} \cdot w_t(\mathcal{R}_t[p_{t+1}]) \right)$.

[6] When a polynomial upper bound n on $\text{OPT}_{\mathcal{F}}$ is known, as per Assumption (A2), we can use it instead of $\text{OPT}_{\mathcal{F}}$ in the formula for T_0; otherwise, exponential search can be used to estimate $\text{OPT}_{\mathcal{F}}$.

Proof.

$$
\begin{aligned}
\Phi(t+1) &= \int_{R\in\mathcal{R}_{t+1}} w_{t+1}(R)dR = \int_{R\in\mathcal{R}_{t+1}} w_t(R)(1-\varepsilon\cdot\mathbb{1}_{p_{t+1}\in R})dR \\
&\le \int_{R\in\mathcal{R}_t} w_t(R)(1-\varepsilon\cdot\mathbb{1}_{p_{t+1}\in R})dR = \Phi(t)\left(1-\varepsilon\int_{R\in\mathcal{R}_t}\mathbb{1}_{p_{t+1}\in R}\frac{w_t(R)}{\Phi(t)}dR\right) \\
&\le \Phi(t)\exp\left(-\varepsilon\int_{R\in\mathcal{R}_t}\mathbb{1}_{p_{t+1}\in R}\frac{w_t(R)}{\Phi(t)}dR\right),
\end{aligned}
$$

where the first inequality is because $\mathcal{R}_{t+1}\subseteq\mathcal{R}_t$ since $|P_t\cap R|$ is non-decreasing in t, and the last inequality is because $1-z\le e^{-z}$ for all z. ◀

▶ **Lemma 8.** *Let* $\kappa(t) := \sum_{t'=0}^{t-1}\frac{w_{t'}(\mathcal{R}_{t'}[p_{t'+1}])}{\Phi(t')}$. *Then* $z^*_{\mathcal{F}}\cdot\kappa(t)\ge\frac{1-\nu}{1+\varepsilon}|P_t|$.

Proof. Due to the choice of $p_{t'+1}$, we have that

$$
\xi_{t'}(p_{t'+1}) := w_{t'}(\mathcal{R}_{t'}[p_{t'+1}]) \ge (1-\nu)\max_{q\in Q} w_{t'}(\mathcal{R}_{t'}[q]). \tag{4}
$$

Consequently, for a $(1+\varepsilon)$-approximate solution μ^*,

$$
\begin{aligned}
z^*_{\mathcal{F}}\cdot\kappa(t) &= \sum_{t'=0}^{t-1} z^*_{\mathcal{F}}\frac{w_{t'}(\mathcal{R}_{t'}[p_{t'+1}])}{\Phi(t')} \ge \frac{1}{1+\varepsilon}\sum_{t'=0}^{t-1}\left(\int_{q\in Q}\mu^*(q)dq\right)\int_{R\in\mathcal{R}_{t'}}\mathbb{1}_{p_{t'+1}\in R}\frac{w_{t'}(R)}{\Phi(t')}dR \\
&\ge \frac{1-\nu}{1+\varepsilon}\sum_{t'=0}^{t-1}\int_{q\in Q}\mu^*(q)\int_{R\in\mathcal{R}_{t'}}\mathbb{1}_{q\in R}\frac{w_{t'}(R)}{\Phi(t')}dRdq \\
&= \frac{1-\nu}{1+\varepsilon}\sum_{t'=0}^{t-1}\int_{R\in\mathcal{R}_{t'}}\left(\int_{q\in Q}\mu^*(q)\mathbb{1}_{q\in R}dq\right)\frac{w_{t'}(R)}{\Phi(t')}dR \\
&= \frac{1-\nu}{1+\varepsilon}\sum_{t'=0}^{t-1}\int_{R\in\mathcal{R}_{t'}}\mu^*(R)\frac{w_{t'}(R)}{\Phi(t')}dR \ge \frac{1-\nu}{1+\varepsilon}\sum_{t'=0}^{t-1}\int_{R\in\mathcal{R}_{t'}}\frac{w_{t'}(R)}{\Phi(t')}dR \\
&= \frac{1-\nu}{1+\varepsilon}\sum_{t'=0}^{t-1}1 = \frac{1-\nu}{1+\varepsilon}|P_t|.
\end{aligned}
$$

where the first inequality is due to the $(1+\varepsilon)$-approximability of μ^*, the second inequality is due to (4), and the last inequality is due to the feasibility of μ^* for (F-HITTING). ◀

▶ **Lemma 9.** *For all* $t=0,1,\dots,$ *we have*

$$
\Phi(t)\le\Phi(0)\exp\left(-\varepsilon\cdot\frac{1-\nu}{1+\varepsilon}\cdot\frac{|P_t|}{z^*_{\mathcal{F}}}\right). \tag{5}
$$

Proof. By repeated application of Lemma 7, and using the result in Lemma 8, we can deduce that

$$
\begin{aligned}
\Phi(t) &\le \Phi(0)\exp\left(-\sum_{t'=0}^{t-1}\frac{\varepsilon}{\Phi(t')}\cdot w_{t'}(\mathcal{R}_{t'}[p_{t'+1}])\right) = \Phi(0)\exp\left(-\varepsilon\kappa(t)\right) \\
&\le \Phi(0)\exp\left(-\varepsilon\cdot\frac{1-\nu}{1+\varepsilon}\cdot\frac{|P_t|}{z^*_{\mathcal{F}}}\right).
\end{aligned}
$$
◀

5.2 Bounding the number of iterations

▶ **Lemma 10.** *After at most* $t_{\max} := \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\nu)}\left(T\ln\frac{1}{1-\varepsilon}+\ln\frac{1}{\varepsilon\delta}\right)$ *iterations, we have* $w_0(\mathcal{R}_{t_{\max}}) < \delta\cdot w_0(\mathcal{R})$.

Proof. For a range $R\in\mathcal{R}$, let us denote by $\mathcal{T}_t(R) := \{0\le t'\le t-1:\ p_{t'+1}\in R\in\mathcal{R}_{t'}\}$ the set of time steps, up to t, at which R was hit by the selected point $p_{t'+1}$, when it was still active. Initialize $w_0'(R) := w_0(R)+\sum_{t'\in\mathcal{T}_t(R)} w_{t'+1}(R)$. For the purpose of the analysis, we will think of the following update step during the algorithm: upon choosing p_{t+1}, set $w_{t+1}'(R) := w_t'(R) - w_t(R)\mathbb{1}_{p_{t+1}\in R}$ for all $R\in\mathcal{R}_t$. Note that the above definition implies that $w_t'(R)\ge(1-\epsilon)^{|\mathcal{T}_t(R)|}w_0(R)$ for all $R\in\mathcal{R}$ and for all t.

▶ **Claim 11.** *For all* t, $w_{t+1}'(\mathcal{R}_{t+1})\le\left(1-\frac{\varepsilon(1-\nu)}{\textsc{Opt}_{\mathcal{F}}}\right)w_t'(\mathcal{R}_t)$.

Proof. Consider an integral optimal solution $P^*\subseteq Q$ (which is guaranteed to exist by (A2)). Then

$$w_t(\mathcal{R}_t) = \int_{R\in\mathcal{R}_t} w_t(R)dR = w_t\left(\bigcup_{q\in P^*}\mathcal{R}_t[q]\right)\le\sum_{q\in P^*}w_t(\mathcal{R}_t[q]). \tag{6}$$

From (6) it follows that there is a $q\in P^*$ such that $w_t(\mathcal{R}_t[q])\ge\frac{w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}$. Note that for such q we have $\xi_t(q) := w_t(\mathcal{R}_t[q])\ge\frac{w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}$, and thus by the choice of p_{t+1}, $\xi_t(p_{t+1})\ge(1-\nu)\xi_t(q)\ge\frac{(1-\nu)w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}$. It follows that

$$\begin{aligned} w_{t+1}'(\mathcal{R}_{t+1}) \le w_{t+1}'(\mathcal{R}_t) &= \int_{R\in\mathcal{R}_t}(w_t'(R)-w_t(R)\mathbb{1}_{p_{t+1}\in R})dR \\ &= \int_{R\in\mathcal{R}_t} w_t'(R)dR - \int_{R\in\mathcal{R}_t} w_t(R)\mathbb{1}_{p_{t+1}\in R}dR = w_t'(\mathcal{R}_t)-\xi_t(p_{t+1}) \\ &\le w_t'(\mathcal{R}_t) - \frac{(1-\nu)w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}. \end{aligned} \tag{7}$$

Note that, for all t,

$$w_t'(R) < w_t(R)\sum_{t'\ge 0}(1-\varepsilon)^{t'} = \frac{w_t(R)}{\varepsilon}. \tag{8}$$

Thus, $w_t(\mathcal{R}_t) > \varepsilon\cdot w_t'(\mathcal{R}_t)$. Using this in (7), we get the claim. ◀

Claim 11 implies that, for $t = t_{\max}$, $w_t'(\mathcal{R}_t)\le\left(1-\frac{\varepsilon(1-\nu)}{\textsc{Opt}_{\mathcal{F}}}\right)^t w_0'(\mathcal{R}) < e^{-\frac{\varepsilon(1-\nu)}{\textsc{Opt}_{\mathcal{F}}}t}w_0'(\mathcal{R})$. Since $|R\cap P_t| < T$ for all $R\in\mathcal{R}_t$, we have $|\mathcal{T}_t(R)|\le T$ and hence $w_t'(\mathcal{R}_t) = \int_{R\in\mathcal{R}_t} w_t'(R)dR\ge(1-\varepsilon)^T w_0(\mathcal{R}_t)$. On the other hand, (8) implies that $w_0'(\mathcal{R}) < \frac{w_0(\mathcal{R})}{\varepsilon}$. Thus, if $w_0(\mathcal{R}_t)\ge\delta\cdot w_0(\mathcal{R})$, we get $(1-\varepsilon)^T\delta < \frac{1}{\varepsilon}\cdot e^{-\frac{\varepsilon(1-\nu)}{\textsc{Opt}_{\mathcal{F}}}t}$, giving $t < \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\nu)}\left(T\ln\frac{1}{1-\varepsilon}+\ln\frac{1}{\varepsilon\delta}\right) = t_{\max}$, in contradiction to $t = t_{\max}$. ◀

5.3 Convergence to an $(\frac{1+5\varepsilon}{1-\nu}, 1-\delta)$-approximate solution

▶ **Lemma 12.** *Suppose that* $T\ge\frac{\max\{1,\ln(g_{\mathcal{F}}(t_{\max})/\delta)\}}{\varepsilon^2}$ *and* $\varepsilon\le 0.67$. *Then Algorithm 1 terminates with a* $(\frac{1+5\varepsilon}{1-\nu}, 1-\delta)$*-approximate solution* $\widehat{\mu}$ *for* (F-hitting).

Proof. Suppose that Algorithm 1 (the while-loop) terminates in iteration $t_f \le t_{\max}$.
$(1-\delta)$-*Feasibility:* By the stopping criterion, $w_0(\mathcal{R}_{t_f}) < \delta \cdot w_0(\mathcal{R})$. Then for $t = t_f$ and any $R \in \mathcal{R} \setminus \mathcal{R}_t$, we have $\widehat{\mu}(R) = \frac{1}{T}\int_{q\in R}\sum_{p\in P_t}\boldsymbol{\delta}_p(q)dq = \frac{1}{T}\sum_{p\in P_t}\int_{q\in R}\boldsymbol{\delta}_p(q)dq = \frac{1}{T}\sum_{p\in P_t}\mathbb{1}_{p\in R} = \frac{1}{T}|P_t \cap R| \ge 1$, since $|P_t \cap R| \ge T$, for all $R \in \mathcal{R}\setminus\mathcal{R}_t$.
Quality of the solution $\widehat{\mu}$*:* We can write

$$\Phi(t) = \sum_{P\in(\mathcal{R}_t)_{|P_t}} (1-\varepsilon)^{|P|} w_0(\mathcal{R}_t[P]), \tag{9}$$

where $\mathcal{R}_t[P] := \{R \in \mathcal{R}_t : \ R\cap P_t = P\}$. Since $\Phi(t)$ satisfies (5), we get by (9) that

$$(1-\varepsilon)^{|P|} w_0(\mathcal{R}_t[P]) \le \Phi(0)\exp\left(-\varepsilon\cdot\frac{1-\nu}{1+\varepsilon}\cdot\frac{|P_t|}{z^*_{\mathcal{F}}}\right), \quad \text{for all } P\in(\mathcal{R}_t)_{|P_t}$$

$$\therefore |P|\ln(1-\varepsilon) + \ln(w_0(\mathcal{R}_t[P])) \le \ln\Phi(0) - \varepsilon\cdot\frac{1-\nu}{1+\varepsilon}\cdot\frac{|P_t|}{z^*_{\mathcal{F}}}, \quad \text{for all } P\in(\mathcal{R}_t)_{|P_t}.$$

Dividing by $\varepsilon\cdot\frac{1-\nu}{1+\varepsilon}\cdot T$ and rearranging, we get

$$\frac{|P_t|}{z^*_{\mathcal{F}}T} \le \frac{(1+\varepsilon)\left(\ln\Phi(0) - \ln(w_0(\mathcal{R}_t[P]))\right)}{\varepsilon(1-\nu)T} + \frac{(1+\varepsilon)|P|}{\varepsilon(1-\nu)T}\cdot\ln\frac{1}{1-\varepsilon}, \quad \text{for all } P\in(\mathcal{R}_t)_{|P_t}. \tag{10}$$

Since $w_0(\mathcal{R}_t) = w_0\left(\bigcup_{P\in(\mathcal{R}_t)_{|P_t}}\mathcal{R}_t[P]\right) = \sum_{P\in(\mathcal{R}_t)_{|P_t}} w_0(\mathcal{R}_t[P])$, there is a set $\widehat{P}\in(\mathcal{R}_t)_{|P_t}$ such that $w_0(\mathcal{R}_t[\widehat{P}]) \ge \frac{w_0(\mathcal{R}_t)}{|(\mathcal{R}_t)_{|P_t}|}$.

We apply (10) for $t = t_f - 1$ and $\widehat{P}\in(\mathcal{R}_t)_{|P_t}$. By the definition of $g_{\mathcal{F}}(\cdot)$, we have $|(\mathcal{R}_t)_{|P_t}| \le g_{\mathcal{F}}(|P_t|) \le g_{\mathcal{F}}(t_{\max})$. Using $\Phi(0) = w_0(\mathcal{R}) \le \frac{w_0(\mathcal{R}_t)}{\delta}$, , $\widehat{\mu}(Q) = \frac{|P_t|+1}{T}$, $|\widehat{P}| < T$ (as $\widehat{P} = R\cap P_t$ for some $R\in\mathcal{R}_t$), $T \ge \frac{\ln(g_{\mathcal{F}}(t_{\max})/\delta)}{\varepsilon^2}$ and $T\ge\frac{1}{\varepsilon^2}$ (by assumption), and $z_{\mathcal{F}^*}\ge 1$, we get (for $\varepsilon\le 0.67$)

$$\begin{aligned}\frac{\widehat{\mu}(Q)}{z^*_{\mathcal{F}}} &\le \frac{(1+\varepsilon)\ln(g_{\mathcal{F}}(t_{\max})/\delta)}{\varepsilon(1-\nu)T} + \frac{(1+\varepsilon)}{\varepsilon(1-\nu)}\cdot\ln\frac{1}{1-\varepsilon} + \frac{1}{T\cdot z^*_{\mathcal{F}}}\\ &\le \frac{\varepsilon(1+\varepsilon)}{(1-\nu)} + \frac{(1+\varepsilon)}{\varepsilon(1-\nu)}\cdot\ln\frac{1}{1-\varepsilon} + \varepsilon^2 < \frac{1+5\varepsilon}{1-\nu}.\end{aligned}$$ ◀

Finally one can verify that the choice of T in (3) satisfies the precondition in Lemma 12.

6 Implementation of the maximization oracle

Let $\mathcal{F} = (Q,\mathcal{R})$ be a range space with VC-dim$(\mathcal{F}) = d$. Recall that the maximization oracle needs to find, for a given $\nu > 0$ and function $w:\mathcal{R}\to\mathbb{R}_+$, a point $p\in Q$ such that $\xi_w(p) \ge (1-\nu)\max_{q\in Q}\xi_w(q)$, where $\xi_w(p) := w(\mathcal{R}[p])$.

To implement the maximization oracle, we follow the approach in [12], based on ϵ-approximations. Recall that an ϵ-approximation for $\mathcal{F}^*$ is a finite subset of ranges $\mathcal{R}'\subseteq\mathcal{R}$, such that (1) holds for all $q\in Q$. We use $\epsilon := \frac{\nu}{2\text{OPT}_{\mathcal{F}}}$ and $\sigma = o(1)$, and take a random sample $\mathcal{R}'$ of size $N = O(\frac{d2^d}{\epsilon^2}\log\frac{1}{\epsilon\sigma}) = O(\frac{d2^d\text{OPT}^2_{\mathcal{F}}}{\nu^2}\log\frac{\text{OPT}_{\mathcal{F}}}{\nu\sigma})$ from $\mathcal{R}$ according to the probability density function $\widehat{w} := w/w(\mathcal{R})$ (for this, we use the sampling oracle $\text{SAMPLE}(\mathcal{F},\widehat{w})$). Then $\mathcal{R}'$ is an ϵ-approximation with high probability. We call $\text{SUBSYS}(\mathcal{F}^*,\mathcal{R}')$ to obtain the set $\mathcal{R}^*_{|\mathcal{R}'}$, then return the subset of ranges $\mathcal{R}''\in\text{argmax}_{\mathcal{R}'''\in\mathcal{R}^*_{|\mathcal{R}'}}|\mathcal{R}'''|$. Finally, we call the oracle $\text{POINTIN}(\mathcal{F},\mathcal{R}'')$ to obtain a point $p\in\bigcap_{R\in\mathcal{R}''}R$.

▶ **Lemma 13.** *With probability* $\Omega(1)$*,* $\xi(p) \geq (1-\nu)\max_{q\in Q}\xi(q)$*.*

▶ Remark. The above implementation of the maximization oracle assumes the *unit-cost* model of computation and *infinite* precision arithmetic (real RAM). In some of the applications in the next section, we note that, in fact, *deterministic* algorithms exist for the maximization oracle, which can be implemented in the *bit-model* with *finite precision*.

7 Applications

7.1 Art gallery problem

In the art gallery problem we are given a (non-simple) polygon H with n vertices and h holes. Two points $p, q \in H$ are said to see each other, denoted by $p \sim q$, if the line segment joining them lies inside H (say, including the boundary ∂H). The objective is to find a subset $G \subseteq H$ such that for every point $q \in H$, there is a point $p \in G$ such that $p \sim q$.

Let $Q = H$, $\mathcal{R} = \{V_H(q) : \ q \in H\}$, where $V_H(q) := \{p \in H : \ p \sim q\}$ is the *visibility region* of $q \in H$. For convenience, we shall consider $\mathcal{R}$ as a *multi-set* and hence assume that ranges in $\mathcal{R}$ are in *one-to-one correspondence* with points in H. We shall see below that the range space $\mathcal{F} = (Q, \mathcal{R})$ satisfies (A1)–(A4).

Note that (A1) is satisfied with $\gamma = \text{VC-dim}(\mathcal{F}) \leq 14$ by the result of [23] for simple polygons, while $\gamma = O(\log h)$ for polygons with h holes [45]. (A2) follows immediately from the fact that each point in the polygon is seen from some vertex. (A3) is satisfied if we use $w_0(R) = 1$ for all $R \in \mathcal{R}$ and note that it is integrable over $\mathcal{R}$ as $\int_{R\in\mathcal{R}} w_0(R)dR = \text{area}(H)$ (recall that ranges in $\mathcal{R}$ are in one-to-one correspondence with points in H). Now we show that (A4) is also satisfied. Consider the randomized implementation of the maximization oracle in Section 6. We need to show that the oracles $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$, $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ and $\textsc{Sample}(\mathcal{F}, w)$ can be implemented in polynomial time. This is more or less standard; we sketch it here for completeness.

It is known (see, e.g., [18]) that the subsystem oracle $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$ can be computed efficiently, for any (finite) $\mathcal{R}' \subset \mathcal{R}$, as follows. Observe that $\mathcal{R}'$ is a finite set of polygons which induces an arrangement of lines (in $\mathbb{R}^2$) of total complexity $O(nh|\mathcal{R}'|^2)$. We can construct the set of cells of this arrangement, call it $\text{cells}(\mathcal{R}')$, in time $O(nh|\mathcal{R}'|^2\log(nh|\mathcal{R}'|))$, and label each cell of the arrangement by the set of visibility polygons from $\mathcal{R}'$ it is contained in. Then $\mathcal{R}^*_{|\mathcal{R}'}$ is the set of different cell labels which can be obtained, for e.g., by a sweep algorithm in time $O(nh|\mathcal{R}'|^2\log(nh|\mathcal{R}'|))$. This argument also implies that $g_{\mathcal{F}^*}(r) \leq O(nhr^2)$, and that we can implement $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ in $O(nh|\mathcal{R}'|^2\log(nh|\mathcal{R}'|))$ time. Finally, we can implement $\textsc{Sample}(\mathcal{F}, \widehat{w}_t)$ given the probability density function $\widehat{w}_t : \mathcal{R} \to \mathbb{R}_+$ defined by the subset $P_t \subseteq Q$ as follows. We construct the cell arrangement $\text{cells}(\mathcal{R})$, induced by the current set P_t as described above. We first sample a cell $\mathcal{R}'$ (which corresponds to an infinite set of ranges with the same weight) with probability $\frac{\widehat{w}_t(\mathcal{R}')}{\sum_{\mathcal{R}'\in\text{cells}(\mathcal{R})}\widehat{w}_t(\mathcal{R}')}$, then we sample a point (corresponding to a range) R uniformly at random from $\mathcal{R}'$. Thus we obtain the following result from Corollary 6, in the *unit-cost* model.

▶ **Corollary 14.** *Given a polygon H with n vertices and h holes and $\delta > 0$, there is a randomized algorithm that finds in* $\text{poly}(n, h, \log\frac{1}{\delta})$ *time a set of points in H of size* $O(z^*_{\mathcal{F}}\log z^*_{\mathcal{F}} \cdot \log(h+2))$ *guarding at least* $(1-\delta)$ *of the area of H, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

We can also obtain a deterministic version of Corollary 14 in the *bit model* of computation. The idea, following [38], is to express the function $\xi_t(q) := w_t(\mathcal{R}_t[q])$ as a sum of continuous functions, each of which is a ratio of two polynomials of two variables, namely, the x and y-coordinates of q. Then maximizing over q amounts to solving a system of two polynomial equations of degree $\text{poly}(n, h, \log \frac{1}{\delta})$ in two variables, which an be solved using *quantifier elimination techniques*, e.g., [5, 25, 41]. However, a technical hurdle that we need to overcome is that the required bit length may grow from one iteration to the next, resulting in an exponential blow-up in the bit length needed for the computation. To deal with this issue, we need to round the set $\mathcal{R}_t$ in each iteration so that the total bit length in all iterations remains bounded by a polynomial in the input size.

▶ **Corollary 15.** *Given a polygon H with n vertices and h holes with rational representation of maximum bit-length L and $\delta > 0$, there is a deterministic algorithm that finds in* $\text{poly}(L, n, h, \log \frac{1}{\delta})$ *time a set of points in H of size $O(z^*_{\mathcal{F}} \log(h+2) \cdot \log(z^*_{\mathcal{F}} \cdot \log(h+2)))$ and bit complexity* $\text{poly}(L, n, h, \log \frac{1}{\delta})$ *guarding at least $(1-\delta)$ of the area of H, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

7.2 Covering a polygonal region by translates of a convex polygon

Let $\mathcal{H}$ be a collection of (non-simple) polygons in the plane and H_0 be a given *full-dimensional convex* polygon. The problem is to minimally cover all the points of the polygons in $\mathcal{H}$ by translates of H_0, that is to find the minimum number of translates $H_0^1, \ldots, H_0^k$ of H_0 such that each point $p \in \bigcup_{H \in \mathcal{H}} H$ is contained in some H_0^i. The discrete case when $\mathcal{H}$ is a set of points has been considered extensively, e.g., covering points with unit disks/squares [28] and generalizations in 3D [15, 32]. Fewer results are known for the continuous case, e.g., [24] which considers the covering of simple polygons by translates of a rectangle[8] and only provides an exact (exponential-time) algorithm; see also [20] for another example, where it is required to hit every polygon in $\mathcal{H}$ by a copy of H_0 (but with rotations allowed).

This problem can be modeled as a hitting set problem in a range space $\mathcal{F} = (Q, \mathcal{R})$, where Q is the set of translates of H_0 and $\mathcal{R} := \left\{\{H_0^i \in Q : \ R \in H_0^i\} : \ R \in \bigcup_{H \in \mathcal{H}} H\right\}$. Again considering $\mathcal{R}$ as a multi-set, we have $\mathcal{R} \leftrightarrow \bigcup_{H \in \mathcal{H}} H$, and we shall refer to elements of $\mathcal{R}$ as sets of translates of H_0 as well as points in $\bigcup_{H \in \mathcal{H}} H$. It was shown by Pach and Woeginger [39] that VC-dim$(\mathcal{F}^*) \le 3$ and also that $\mathcal{F}^*$ admits an ϵ-net of size $s_{\mathcal{F}^*} = O(\frac{1}{\epsilon})$. As observed in [32], this would also imply that VC-dim$(\mathcal{F}) \le 3$ and $s_{\mathcal{F}} = O(\frac{1}{\epsilon})$. Thus (A1) is satisfied with $\gamma = 3$. Moreover, assuming that $\mathcal{H}$ is contained in a box of size D and that H_0 contains a box of size d, then (A2) is satisfied as $\textsc{Opt}_{\mathcal{F}} \le \frac{D}{d}$. (A3) is satisfied if we use $w_0(R) = 1$ for all $R \in \mathcal{R}$ (which defines the area measure over $\mathcal{R}$). Now we show that (A4) is also satisfied.

Consider the randomized implementation of the maximization oracle in Section 6. We need to show that the oracles $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$, $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ and $\textsc{Sample}(\mathcal{F}, w)$ can be implemented in polynomial time. Let m be the total number of vertices of the polygons in $\mathcal{H}$ and H_0. Note that for a given finite $\mathcal{R}' \subseteq \mathcal{R}$, the set $\mathcal{R}^*_{|\mathcal{R}'}$ is the set of all subsets of points in $\mathcal{R}'$ that are contained in the same copy of H_0. Observe that each such subset is determined by at most two points from $\mathcal{R}'$ that lie on the boundary of a copy of H_0. It follows that $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$ can be implemented in $O(m^2|\mathcal{R}'|^2 \log m)$ time. This argument also shows that $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ can be implemented in the same time $O(m^2|\mathcal{R}'|^3)$ and that $g_{\mathcal{F}^*}(r) \le r^2$. Finally, we can implement $\textsc{Sample}(\mathcal{F}, \widehat{w}_t)$ given the probability density function $\widehat{w}_t : \mathcal{R} \to \mathbb{R}_+$ defined by the subset $P_t \subseteq Q$ as follows. Given the current subset $P_t \subseteq Q$ of translates of H_0, we can find (e.g. by a sweep line algorithm) in $O(m \log m)$ time the cells of the arrangement defined by $\mathcal{H} \cup P_t$ (where a cell is naturally defined to be a

[8] Note that in [24], each polygon has to be covered *completely* by a rectangle.

maximal set of points in $\mathcal{R}$ that all belong exactly to the same polygons in the arrangement). Let us call this set cells$(\mathcal{R})$ and note that it has size $O(m)$. We first sample a cell $\mathcal{R}'$ with probability $\frac{\widehat{w}_t(\mathcal{R}')}{\sum_{\mathcal{R}'\in \text{cells}(\mathcal{R})}\widehat{w}_t(\mathcal{R}')}$, then we sample a point R uniformly at random from $\mathcal{R}'$.

▶ **Corollary 16.** *Given a collection of polygons in the plane $\mathcal{H}$ and a convex polygon H_0, with m total vertices and $\delta > 0$, there is a randomized algorithm that finds in* $\text{poly}(n, m, \log \frac{1}{\delta})$ *time a set of $O(z^*_{\mathcal{F}})$ translates of H_0 covering at least $(1-\delta)$ of the total area of the polygons in $\mathcal{H}$, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

7.3 Polyhedral separation in $\mathbb{R}^d$

Given two (full-dimensional) convex polytopes $\mathcal{P}_1, \mathcal{P}_2 \subseteq \mathbb{R}^d$ such that $\mathcal{P}_1 \subset \mathcal{P}_2$, it is required to find a (separator) polytope $\mathcal{P}_3 \subseteq \mathbb{R}^d$ such that $\mathcal{P}_1 \subseteq \mathcal{P}_3 \subseteq \mathcal{P}_2$, with as few facets as possible. This problem can be modeled as a hitting set problem in a range space $\mathcal{F} = (Q, \mathcal{R})$, where Q is the set of supporting hyperplanes for $\mathcal{P}_1$ and $\mathcal{R} := \{\{p \in Q : p \text{ separates } R \text{ from } \mathcal{P}_1\} : R \in \partial\mathcal{P}_2\}$ (thus, we may assume that $\mathcal{R} \leftrightarrow \partial\mathcal{P}_2$). Note that VC-dim$(\mathcal{F}) = d$ (and VC-dim$(\mathcal{F}^*) = d+1$). In their paper [8], Brönnimann and Goodrich gave a deterministic $O(d^2 \log \textsc{Opt}_{\mathcal{F}})$-approximation algorithm, improving on earlier results by Mitchell and Suri [37], and Clarkson [14]. It was shown in [37] that, at the cost of losing a factor of d in the approximation ratio, one can consider a finite set Q, consisting of the hyperplances passing through the facets of $\mathcal{P}_1$. We can save this factor of d by showing that $\mathcal{F}$ satisfies (A1)–(A4).

Let n and m be the number of facets of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Then (A1) is satisfied with $\gamma = d$ as explained above. Also, (A2) is obviously satisfied since $\mathcal{P}_3 = \mathcal{P}_2$ is a separator with n facets. For (A3), we use w_0 as the *surface area* measure, i.e., $w_0(\mathcal{R}') = \text{vol}_{d-1}(\mathcal{R}')$ for $\mathcal{R}' \subseteq \mathcal{R}$. Now we show that (A4) also holds.

Consider the randomized implementation of the maximization oracle in Section 6. We need to show that the oracles $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$, $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ and $\textsc{Sample}(\mathcal{F}, w)$ can be implemented in polynomial time. Note that for a given finite $\mathcal{R}' \subseteq \mathcal{R}$, the set $\mathcal{R}^*_{|\mathcal{R}'}$ has size at most $g(|\mathcal{R}'|, d+1)$, and furthermore, for any hyperplane $q \in Q$, $\mathcal{R}'[q]$ is the set of points in $\mathcal{R}'$ separated from $\mathcal{P}_1$ by q. Thus, $\mathcal{R}'[q]$ is determined by exactly d points chosen from $\mathcal{R}'$ and the vertices of $\mathcal{P}_1$. It follows that the set $\mathcal{R}^*_{|\mathcal{R}'}$ can be found (and hence $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$ can be implemented) in time $\text{poly}((n^{\frac{d}{2}} + |\mathcal{R}'|)^d)$. This argument also shows that $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ can be implemented in the time $\text{poly}((n^{\frac{d}{2}} + |\mathcal{R}'|)^d)$. Finally, we can implement $\textsc{Sample}(\mathcal{F}, \widehat{w}_t)$ given the probability density function $\widehat{w}_t : \mathcal{R} \to \mathbb{R}_+$ defined by the current subset $P_t \subseteq Q$ as follows. We first construct the set of cells of the hyperplane arrangement of P_t, which has complexity $O(|P_t|^d)$, in time $O(|P_t|^{d+1})$; see, e.g., [4, 44]. Next, we intersect every facet of $\mathcal{P}_2$ with every cell in the arrangement. This allows us to identify the partition of $\partial\mathcal{P}_2$ induced by the cell arrangement; let us call it cells$(\mathcal{R})$. The running time for this is $\text{poly}(|P_t|^d, m^d)$. We first sample $\mathcal{R}'$ with probability $\frac{\widehat{w}_t(\mathcal{R}')}{\sum_{\mathcal{R}'\in \text{cells}(\mathcal{R})}\widehat{w}_t(\mathcal{R}')}$, then we sample a point R uniformly at random from $\mathcal{R}'$ (note that both volume computation and uniform sampling can be done in polynomial time in fixed dimension).

▶ **Corollary 17.** *Given two convex polytopes $\mathcal{P}_1, \mathcal{P}_2 \subseteq \mathbb{R}^d$ such that $\mathcal{P}_1 \subset \mathcal{P}_2$, with n and m facets respectively and $\delta > 0$, there is a randomized algorithm that finds in* $\text{poly}((nm)^d, \log \frac{1}{\delta})$ *time a polytope $\mathcal{P}_3$ with $O(z^*_{\mathcal{F}} d \cdot \log z^*_{\mathcal{F}})$ facets separating $\mathcal{P}_1$ from a subset of $\partial\mathcal{P}_2$ of volume at least $(1-\delta)$ of the volume of $\partial\mathcal{P}_2$, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

Acknowledgement. The author is grateful to Waleed Najy for his help in the proof of Lemma 12 and for many useful discussions, and to the anonymous reviewers for the careful reading and the helpful remarks.

References

1 P. K. Agarwal and J. Pan. Near-linear algorithms for geometric hitting sets and set covers. In *SoCG'14*, pages 271–279, 2014.

2 N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2008.

3 B. Aronov, E. Ezra, and M. Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM Journal on Computing*, 39(7):3248–3282, 2010.

4 D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete & Computational Geometry*, 8(3):295–313, 1992.

5 S. Basu, R. Pollack, and M. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *J. ACM*, 43(6):1002–1045, 1996.

6 É. Bonnet and T. Miltzow. An approximation algorithm for the art gallery problem. In *EuroCG'16, also available online as: https://arxiv.org/abs/1607.05527*, 2016.

7 H. Brönnimann, B. Chazelle, and J. Matoušek. Product range spaces, sensitive sampling, and derandomization. *SIAM Journal on Computing*, 28(5):1552–1575, 1999.

8 H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete & Computational Geometry*, 14(4):463–479, 1995.

9 T. M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *SODA'12*, pages 1576–1585, 2012.

10 B. Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, New York, NY, USA, 2000.

11 B. Chazelle and J.Matoušek. On linear-time deterministic algorithms for optimization problems in fixed dimension. *Journal of Algorithms*, 21(3):579–597, 1996.

12 O. Cheong, A. Efrat, and S. Har-Peled. Finding a guard that sees most and a shop that sells most. *Discrete & Computational Geometry*, 37(4):545–563, 2007.

13 V. Chvatal. A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, 4(3):233–235, 1979.

14 K. L. Clarkson. Algorithms for polytope covering and approximation. In *WADS'93*, pages 246–252, 1993.

15 K. L. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete & Computational Geometry*, 37(1):43–58, 2006.

16 A. Deshpande, T. Kim, E. D. Demaine, and S. E. Sarma. A pseudopolynomial time $O(\log n)$-approximation algorithm for art gallery problems. In *WADS'07*, pages 163–174, 2007.

17 I. Dinur and D. Steurer. Analytical approach to parallel repetition. In *STOC'14*, pages 624–633, 2014.

18 A. Efrat and S. Har-Peled. Guarding galleries and terrains. *Inf. Process. Lett.*, 100(6):238–245, 2006.

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

20 S. K. Ganjugunte. *Geometric Hitting Sets and Their Variants*. PhD thesis, Duke University, USA, 2011.

21 N. Garg and J. Könemann. Faster and simpler algorithms for multicommodity flow and other fractional packing problems. *SIAM J. Comput.*, 37(2):630–652, 2007.

22 S. K. Ghosh. Approximation algorithms for art gallery problems in polygons. *Discrete Applied Mathematics*, 158(6):718–722, 2010.

23 A. Gilbers and R. Klein. A new upper bound for the VC-dimension of visibility regions. *Computational Geometry*, 47(1):61–74, 2014.

24 R. Glück. Covering polygons with rectangles. In *EuroCG'16*, 2016.

25 D. Grigoriev and N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, 5(1/2):37–64, 1988.

26 S. Har-Peled and M. Sharir. Relative (p, ϵ)-approximations in geometry. *Discrete & Computational Geometry*, 45(3):462–496, 2011.

27 D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.

28 D. S. Hochbaum and W. Maass. Approximation schemes for covering and packing problems in image processing and VLSI. *J. ACM*, 32(1):130–136, 1985.

29 D. S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3):256–278, 1974.

30 J. King and D. G. Kirkpatrick. Improved approximation for guarding simple galleries from the perimeter. *Discrete & Computational Geometry*, 46(2):252–269, 2011.

31 J. Komlós, J. Pach, and G. Woeginger. Almost tight bounds for ϵ-nets. *Discrete & Computational Geometry*, 7(2):163–173, March 1992.

32 S. Laue. Geometric set cover and hitting sets for polytopes in $\mathbb{R}^3$. In *STACS'08*, pages 479–490, 2008.

33 L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13(4):383–390, 1975.

34 J. Matoušek. Cutting hyperplane arrangements. *Discrete & Computational Geometry*, 6(3):385–406, 1991.

35 J. Matoušek. Reporting points in halfspaces. *Computational Geometry*, 2(3):169–186, 1992.

36 J. Matoušek, R. Seidel, and E. Welzl. How to net a lot with little: Small ϵ-nets for disks and halfspaces. In *SoCG'90*, pages 16–22, 1990.

37 J. S. B. Mitchell and S. Suri. Separation and approximation of polyhedral objects. *Computational Geometry*, 5(2):95–114, 1995.

38 S. Ntafos and M. Tsoukalas. Optimum placement of guards. *Information Sciences*, 76(1–2):141–150, 1994.

39 J. Pach and G. Woeginger. Some new bounds for Epsilon-nets. In *SoCG'90*, pages 10–15, 1990.

40 E. Pyrga and S. Ray. New existence proofs ϵ-nets. In *SoCG'08*, pages 199–207, 2008.

41 J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. *J. Symb. Comput.*, 13(3):255–352, 1992.

42 N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

43 S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41(1):247–261, 1972.

44 N. H. Sleumer. Output-sensitive cell enumeration in hyperplane arrangements. *Nordic J. of Computing*, 6(2):137–147, 1999.

45 P. Valtr. Guarding galleries where no point sees a small area. *Israel Journal of Mathematics*, 104(1):1–16, 1998.

46 V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971.

47 K. Varadarajan. Epsilon nets and union complexity. In *SoCG'09*, pages 11–16, 2009.