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

Khaled Elbassioni*

October 13, 2016

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. These 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 polylog$(\frac{1}{\delta})$.

1 Introduction

Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space defined by a 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 [31, 37, 14] yield $(\log |Q| + 1)$-approximation in polynomial time, and this is essentially the best possible guarantee assuming $NP \not\subset DTIME(n^{O(\log\log n)})$ [21, 38]. Better approximation algorithms exist for special cases, such as range spaces of *bounded VC-dimension* [7], of *bounded union complexity* [16, 52], of *bounded shallow cell complexity* [9], as well as several classes of geometric range spaces [3, 45, 34]. 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 weight updates algorithm of Brönnimann and Goodrich [7]. For instance, if a range space $\mathcal{F}$ has VC-dimension d then it admits an ϵ-net of size $O(\frac{d}{\epsilon}\log\frac{1}{\epsilon})$ [29, 35], which by the above mentioned method implies an $O(d \cdot \log \textsc{Opt}_{\mathcal{F}})$-approximation algorithm for the hitting set problem for $\mathcal{F}$, where $\textsc{Opt}_{\mathcal{F}}$ denotes the size of a minimum-size hitting set. Even et al. [20] 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[1] $z^*_{\mathcal{F}}$, and then finding an ϵ-net, with $\epsilon := 1/z^*_{\mathcal{F}}$.

The multiplicative weight updates algorithm in [7] 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

*Masdar Institute of Science and Technology, P.O. Box 54224, Abu Dhabi, UAE; (kelbassioni@masdar.ac.ae)

[1] In fact, we will observe below (see Appendix A) that the exact same algorithm of [7], but with a slightly modified analysis, gives this improved bound of [20], *without* the need to solve an LP.

be arbitrarily small (see Appendix A for details). In this paper we take a different approach, which can be thought of as a combination of the methods in [7] and [20] (with LP replaced by an *infinite dimensional convex relaxation*):

- We maintain weights on the *ranges* (in contrast to Brönnimann and Goodrich [7] which maintain 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 multiplicative weight updates (MWU), extending the approach in [23] to infinite dimensional covering LP's (under reasonable assumptions);

- We finally use the rounding idea of [20] to get a small integral hitting set from the obtained fractional solution.

Informal main theorem. Given a range space $\mathcal{F}=(Q,\mathcal{R})$ of VC-dimension d, (under mild assumptions) there is an algorithm that, for any $\delta>0$, 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 $\log(\frac{1}{\delta})$.

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 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 [13] which gives a polytime algorithm that finds a set of size $O(\textsc{Opt}_{\mathcal{F}}\cdot\log\frac{1}{\delta})$ hitting $(1-\delta)$-fraction of $\mathcal{R}$. 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 [17], restrict the set of candidate guard locations [18], or make some general position assumptions [6].

- Covering a polygonal region by translates of a convex polygon: Given a collection of polygons in the plane $\mathcal{H}$ 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 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 [16, 30, 36].

- 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 [7] (but which gives a complete separation).

More related work on these problems can be found in the corresponding subsections of Section 7.

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*; this will be explained in Section 7.1. 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. The *dual* range space $\mathcal{F}^* = (Q^*, \mathcal{R}^*)$ is defined as the range space with $Q^* := \mathcal{R}$ and $\mathcal{R}^* := \{\{R \in \mathcal{R} : \ q \in \mathcal{R}\} : \ q \in Q\}$. For a point $q \in Q$ and a subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$, let $\mathcal{R}'[q] := \{R \in \mathcal{R}' : \ q \in R\}$. 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 $\mathcal{Q}_{\mathcal{R}'} := \{\mathcal{R}'[q] : \ q \in Q\}$. For a finite set $P \subseteq Q$ of size r, we denote by $g_{\mathcal{F}}(r) \leq 2^r$ the smallest integer such that $|\mathcal{R}_P| \leq g_{\mathcal{F}}(r)$. 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 Problem definition and assumptions

More formally, we consider the following problem:

Min-Hitting-Set: Given a range space $\mathcal{F} = (Q, \mathcal{R})$, find a minimum-size hitting set.

We shall make the following assumptions[2]:

(A1) $g_{\mathcal{F}}(r) \leq r^{\gamma}$, for some non-decreasing function $g : \mathbb{N} \to \mathbb{R}_+$, and some constant $\gamma \geq 1$.

(A1′) The range space is given by a *subsystem oracle* Subsys$(\mathcal{F}, P)$ that, given any finite $P \subseteq Q$, returns the set of ranges $\mathcal{R}|_P$.

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

(A3) There exists a finite measure $w_0 : \mathcal{R} \to \mathbb{R}_+$ such that all subsets of $\mathcal{R}$ are w_0-measurable.

2.3 Range spaces of bounded VC-dimension

We consider range spaces 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 the VC-dim$(\mathcal{F}) = +\infty$. It is well-known that if VC-dim$(\mathcal{F}) = d$ then $g_{\mathcal{F}}(r) \leq O(r^d)$. More precisely, the following bound holds.

Lemma 1 (Sauer-Shelah Lemma [47, 48])**.** *For any range space $\mathcal{F} = (Q, \mathcal{R})$ of VC-dimension d and any $r \geq 1$, it holds that $g_{\mathcal{F}}(r) \leq g(r, d) := \sum_{i=0}^{d} \binom{r}{i}$.*

Lemma 2. *If VC-dim$(\mathcal{F}) = d$ then VC-dim$(\mathcal{F}^*) < 2^{d+1}$.*

[2] For simplicity of presentation, we will make the implicit assumption in this paper that both Q and $\mathcal{R}$ is are in one-to-one correspondence with some subsets of $\mathbb{R}^k$, as all the applications we consider have this restriction. This implies that the measure w_0 in (A3) (and μ_0 in (A3′)) can be taken as the standard volume measures in $\mathbb{R}^k$, and the integrals used below are the standard Riemann integrals. However, we note that the extension to general measurable sets should be straightforward.

2.4 ϵ-nets

Given a range space $(Q, \mathcal{R})$, a finite measure $\mu : Q \to \mathbb{R}_+$ (such that the ranges in $\mathcal{R}$ are μ-measurable), 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 $\mu(R) \geq \epsilon \cdot \mu(Q)$. We say that a range space $\mathcal{F}$ admits an ϵ-net of size $s_{\mathcal{F}}(\cdot)$, if for any $\epsilon > 0$, there an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$. For range spaces of VC-dimension d, Haussler and Welzl [29] proved a bound of $s_{\mathcal{F}}(\frac{1}{\epsilon}) := O(\frac{d}{\epsilon} \log \frac{d}{\epsilon})$ on the size of an ϵ-net, which was later slightly improved by Komlós et al.

Theorem 3 (ϵ-net Theorem [29, 35])**.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d, μ be an arbitrary probability measure on Q (such that the ranges in $\mathcal{R}$ are μ-measurable), and $\epsilon > 0$ be a given parameter. Then there exists an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon}) = O(\frac{d}{\epsilon} \log \frac{1}{\epsilon})$. In fact, a random sample (w.r.t. to the probability measure μ) of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$ is an ϵ-net with (high) probability $\Omega(1)$.*

We say that a finite measure $\mu : Q \to \mathbb{R}_+$ has (finite) support K if μ can be written as a conic combination of K *Dirac measures*[3]: $\mu = \sum_{p \in P} \mu(p) \boldsymbol{\delta}_p(q)$, for some finite $P \subseteq Q$ of cardinality K and non-negative multipliers $\mu(p)$, for $p \in P$. Measures of finite support can be considered as weights on a finite subset of Q, in which case an ϵ-net can be computed deterministically as given by the following result of Matoušek [39].

Theorem 4 ([8, 11, 39])**.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d satisfying (A1′), μ be a measure on Q with support K, and $\epsilon > 0$ be a given a parameter. Then for any $\epsilon > 0$, there is a deterministic algorithm that computes an ϵ-net for $\mathcal{R}$ of size $s_{\mathcal{F}}(\frac{1}{\epsilon}) = O(\frac{d}{\epsilon} \log \frac{d}{\epsilon})$ in time $O(d)^{3d} \frac{1}{\epsilon^{2d}} \log^d(\frac{d}{\epsilon}) K$.*

Since most of the results on ϵ-nets are stated in terms of the unweighted case, it is worth recalling the reduction from the weighted case to the unweighted case (see , e.g., [39]). Given a measure μ defined on a finite set P of support K, we replace each point $p \in P$, by $\left\lfloor \frac{\mu(p)K}{\sum_{p \in P} \mu(p)} + 1 \right\rfloor$ copies of p. Let Q' be the new set of points. Then $K' := |Q'| \leq 2K$ and an $\frac{\epsilon}{2}$-net for $(Q', \mathcal{R}|_{Q'})$ is an ϵ-net for $(Q, \mathcal{R}|_Q)$.

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 $\mathbb{R}^3$ [41, 40]; see also [9, 34, 32, 52].

2.5 ϵ-approximations

Given the dual range space $\mathcal{F}^*$, a measure $w : \mathcal{R} \to \mathbb{R}_+$, and an $\epsilon > 0$, an ϵ-approximation is a finite subset of ranges $\mathcal{R}' \subseteq \mathcal{R}$ such that, for all $q \in Q$,

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

see, e.g., [10]. The following theorem (stated in the dual space for our purposes, where VC-dim$(\mathcal{F}^*) < 2^{d+1}$ by Lemma 2) states the existence of an ϵ-approximation of small size.

Theorem 5 (ϵ-approximation Theorem [2, 10, 51])**.** *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d, w be an arbitrary probability measure on $\mathcal{R}$, and $\epsilon > 0$ be a given a parameter. Then a random sample (w.r.t. the probability measure w) of size $O(\frac{d2^d}{\epsilon^2} \log \frac{1}{\epsilon\sigma})$ is an ϵ-approximation for $\mathcal{F}^*$, with probability $1 - \sigma$.*

[3] The Dirac measure satisfies $\int_{Q'} \boldsymbol{\delta}_p(q) dq = 1$ if $p \in Q'$, and $\int_{Q'} \boldsymbol{\delta}_p(q) dq = 0$ otherwise.

2.6 The fractional problem

Given a range space $\mathcal{F} = (Q, \mathcal{R})$, satisfying assumptions (A1)-(A3), the fractional problem seeks to find a measure μ on Q, such that $\mu(R) \geq 1$ for all $R \in \mathcal{R}$ and $\mu(Q)$ is minimized[4]:

$$\begin{aligned} z^*_{\mathcal{F}} := \inf_{\mu} & \int_{q \in Q} \mu(q) dq && \text{(F-HITTING)} \\ \text{s.t.} & \int_{q \in R} \mu(q) dq \geq 1, \forall R \in \mathcal{R}, && (2) \\ & \mu(q) \geq 0, \forall q \in Q. \end{aligned}$$

Equivalently, it is required to find a *probability* measure $\mu : Q \to [0, 1]$ that solves the maximin problem: $\sup_{\mu} \inf_{R \in \mathcal{R}} \mu(R)$.

Proposition 6. *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 a measure μ of support $\text{OPT}_{\mathcal{F}}$ by $\mu(q) := \sum_{p \in P^*} \boldsymbol{\delta}_p(q)$. 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 α-approximate and β-feasible.

2.7 Rounding the fractional solution

Brönnimann and Goodrich [7] gave a multiplicative weight updates algorithm for approximating the minimum hitting set for a *finite* range space satisfying (A1′) and admitting an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$. For completeness, their algorithm is given as Algorithm 2 in Appendix A, and 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 invokes Theorem 4 to find 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. [20] 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 7. *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 1. *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space of VC-dimension d and μ be a measure on Q satisfying (2). Then a random sample of size $O(d \cdot \mu(Q) \log(\mu(Q)))$, w.r.t. the probability measure $\mu' := \frac{\mu}{\mu(Q)}$, is a hitting set for $\mathcal{R}$ with probability $\Omega(1)$. Furthermore, if μ has support 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 $O(d)^{3d} \mu(Q)^{2d} \log^d(d \cdot \mu(Q))K$.*

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

Proof. In view of Lemma 7, the two parts of the corollary follow from Theorems 3 and 4, respectively. □

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

3 Solving the fractional problem – Main result

We make the following further assumption:

(A4) There is a deterministic (resp., randomized) oracle $\textsc{Max}(\mathcal{F}, w, \omega)$ (resp., $\textsc{Max}(\mathcal{F}, w, \sigma, \omega)$), that given a range space $\mathcal{F} = (Q, \mathcal{R})$, a finite measure $w : \mathcal{R} \to \mathbb{R}_+$ on $\mathcal{R}$, and $\omega > 0$, returns (resp., with probability $1 - \sigma$) a point $p \in Q$ such that

$$\xi_w(p) \geq (1 - \omega) \max_{q \in Q} \xi_w(q),$$

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

In view of Corollary 1, we have the following theorem as an immediate consequence of Theorem 8.

Theorem 9 (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 $\varepsilon, \delta, \omega \in (0, 1)$ 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 $O(\frac{\gamma}{\varepsilon^3(1-\omega)} \log \frac{\gamma}{\varepsilon} \cdot \textsc{Opt}_{\mathcal{F}} \log \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\omega)})$ calls to the oracle $\textsc{Max}(\ldots, \omega)$ 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 can extend Corollary 1 as follows (under the assumption of the availability of subsystem and sampling oracles in the dual range space); see Section 6 for details.

Corollary 2. *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 $O(d \cdot z^*_{\mathcal{F}} \log(d \cdot z^*_{\mathcal{F}}))$, hitting a subset of $\mathcal{R}$ of measure at least $(1 - \delta)w_0(\mathcal{R})$, in time $O(K \cdot g_{\mathcal{F}^*}(\frac{d2^d\textsc{Opt}^2_{\mathcal{F}}}{\varepsilon^2} \log \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon}))$, where $K := O(\frac{d}{\varepsilon^3(1-\varepsilon)} \log \frac{d}{\varepsilon} \cdot \textsc{Opt}_{\mathcal{F}} \log \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\varepsilon)})$.*

Note by Lemma 2 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

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

Clearly, (since these properties are hereditary) VC-dim($\mathcal{F}_t$) $\le$ VC-dim($\mathcal{F}$), and $\mathcal{F}_t$ admits and 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

$$T_0 := \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\omega)\delta^{1/\gamma}}\left(\ln\frac{1}{1-\varepsilon}+\ln\frac{1}{\varepsilon\delta}\right), \quad a := \frac{\gamma}{\varepsilon^2}, \quad \text{and } b := \max\{\ln T_0, 1\},$$

$$T := e^2ab(\ln(a+e-1)+1) = \Theta(\frac{\gamma}{\varepsilon^2}\log\frac{\gamma}{\varepsilon}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\omega)}). \tag{3}$$

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

```
  Data: A range space $\mathcal{F} = (Q, \mathcal{R})$ satsfying (A1)-(A4), and an approximation accuracies
        $\varepsilon, \delta, \omega, \in (0, 1)$.
  Result: A $(\frac{1+5\varepsilon}{1-\omega}, 1-\delta)$-approximate solution $\mu$ for (F-HITTING).
1 $t \leftarrow 0$; $P_0 \leftarrow \emptyset$; set $T$ as in (3)
2 while $w_0(\mathcal{R}_t) \ge \delta \cdot w_0(\mathcal{R})$ do
3     define the measure $w_t : \mathcal{R}_t \to \mathbb{R}_+$ by $w_t(R) \leftarrow (1-\varepsilon)^{|P_t \cap R|} w_0(R)$, for $R \in \mathcal{R}_t$
4     $p_{t+1} \leftarrow \textsc{Max}(\mathcal{F}_t, w_t, \omega)$
5     $P_{t+1} \leftarrow P_t \cup \{p_{t+1}\}$
6     $t \leftarrow t+1$
7 end
8 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

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 4 upto 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 10. *For all $t = 0, 1, \ldots$, it holds that*

$$\Phi(t+1) \le \Phi(t)\exp\left(-\frac{\varepsilon}{\Phi(t)} \cdot w_t(\mathcal{R}_t[p_{t+1}])\right). \tag{4}$$

Proof.

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

$$\leq \Phi(t) \exp\left(-\varepsilon \int_{R\in\mathcal{R}_t} \mathbb{1}_{p_{t+1}\in R} \frac{w_t(R)}{\Phi(t)} dR\right),$$

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 $1 - z \leq e^{-z}$ for all z. □

Lemma 11. *Let* $\kappa(t) := \sum_{t'=0}^{t-1} \frac{w_{t'}(\mathcal{R}_{t'}[p_{t'+1}])}{\Phi(t')}$. *Then* $z^*_{\mathcal{F}} \cdot \kappa(t) \geq \frac{1-\omega}{1+\varepsilon}|P(t)|$.

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

$$\xi_{t'}(p_{t'+1}) := w_{t'}(\mathcal{R}_{t'}[p_{t'+1}]) \geq (1-\omega) \max_{q\in Q} w_{t'}(\mathcal{R}_{t'}[q]). \tag{5}$$

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

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

where the first inequality is due to the $(1+\varepsilon)$-optimality of μ^*, the second inequality is due to (5), and the last inequality is due to the feasibility of μ^* for (F-HITTING). □

Lemma 12. *For all* $t = 0, 1, \ldots,$ *we have*

$$\Phi(t) \leq \Phi(0) \exp\left(-\varepsilon \cdot \frac{1-\omega}{1+\varepsilon} \cdot \frac{|P(t)|}{z^*_{\mathcal{F}}}\right). \tag{6}$$

Proof. By repeated application of (4), and using the result in Lemma 11, we can deduce that

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

□

5.2 Bounding the number of iterations

Lemma 13. *After at most* $t_{\max} := \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\omega)}\left(T \ln \frac{1}{1-\varepsilon} + \ln \frac{1}{\varepsilon\delta}\right)$ *iterations, we have* $w_0(\mathcal{R}_{t_f}) < \delta \cdot w_0(\mathcal{R})$.

Proof. For a range $R \in \mathcal{R}$, let us denote by $\mathcal{T}_t(R) := \{0 \leq t' \leq t-1 : \ p_{t'+1} \in R \in \mathcal{R}_{t'}\}$ the set of time steps, upto 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 $w'_t(R) \geq (1-\epsilon)^{|\mathcal{T}_t(R)|} w_0(R)$ for all $R \in \mathcal{R}$ and for all t.

Claim 14. *For all t:*

$$w'_{t+1}(\mathcal{R}_{t+1}) \leq \left(1 - \frac{\varepsilon(1-\omega)}{\textsc{Opt}_{\mathcal{F}}}\right) w'_t(\mathcal{R}_t). \tag{7}$$

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

$$w_t(\mathcal{R}_t) = \int_{R\in\mathcal{R}_t} w_t(R)dR = w_t\left(\bigcup_{q\in P^*} \mathcal{R}_t[q]\right) \leq \sum_{q\in P^*} w_t(\mathcal{R}_t[q]). \tag{8}$$

From (8) it follows that there is a $q \in P^*$ such that $w_t(\mathcal{R}_t[q]) \geq \frac{w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}$. Note that for such q we have

$$\xi_t(q) := w_t\left(\mathcal{R}_t[q]\right) \geq \frac{w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}, \tag{9}$$

and thus by the choice of p_{t+1}, $\xi_t(p_{t+1}) \geq (1-\omega)\xi_t(q) \geq \frac{(1-\omega)w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}$. It follows that

$$\begin{aligned} w'_{t+1}(\mathcal{R}_{t+1}) \leq w'_{t+1}(\mathcal{R}_t) &= \int_{R\in\mathcal{R}_t} (w'_t(R) - w_t(R)\mathbb{1}_{p_{t+1}\in R})dR \\ &= \int_{R\in\mathcal{R}_t} w'_t(R)dR - \int_{R\in\mathcal{R}_t} w_t(R)\mathbb{1}_{p_{t+1}\in R}dR = w'_t(\mathcal{R}_t) - \xi_t(p_{t+1}) \\ &\leq w'_t(\mathcal{R}_t) - \frac{(1-\omega)w_t(\mathcal{R}_t)}{\textsc{Opt}_{\mathcal{F}}}. \end{aligned} \tag{10}$$

Note that, for all t,

$$w'_t(R) < w_t(R)\sum_{t'\geq 0}(1-\varepsilon)^{t'} = \frac{w_t(R)}{\varepsilon}. \tag{11}$$

Thus, $w_t(\mathcal{R}_t) > \varepsilon \cdot w'_t(\mathcal{R}_t)$. Using this in (10), we get the claim. □

Claim 14 implies that, for $t = t_{\max}$,

$$w'_t(\mathcal{R}_t) \leq \left(1 - \frac{\varepsilon(1-\omega)}{\textsc{Opt}_{\mathcal{F}}}\right)^t w'_0(\mathcal{R}) < e^{-\frac{\varepsilon(1-\omega)}{\textsc{Opt}_{\mathcal{F}}}t} w'_0(\mathcal{R}).$$

Since $|R \cap P_t| < T$ for all $R \in \mathcal{R}_t$, we have $w'_t(\mathcal{R}_t) = \int_{R\in\mathcal{R}_t} w'_t(R)dR > (1-\varepsilon)^T w_0(\mathcal{R}_t)$. On the other hand, (11) implies that $w'_0(\mathcal{R}) < \frac{w_0(\mathcal{R})}{\varepsilon}$. Thus, if $w_0(\mathcal{R}_t) \geq \delta \cdot w_0(\mathcal{R})$, we get

$$(1-\varepsilon)^T\delta < \frac{1}{\varepsilon} \cdot e^{-\frac{\varepsilon(1-\omega)}{\textsc{Opt}_{\mathcal{F}}}t},$$

giving $t < \frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\omega)}\left(T \ln \frac{1}{1-\varepsilon} + \ln \frac{1}{\varepsilon\delta}\right) = t_{\max}$, in contradiction to $t = t_{\max}$. □

5.3 Convergence to an $(\frac{1+4\varepsilon}{1-\omega}, 1-\delta)$-approximate solution

Lemma 15. *Suppose that $T \geq \frac{\max\{1, \ln(g_{\mathcal{F}}(t_{\max})/\delta)\}}{\varepsilon^2}$ and $\varepsilon \leq 0.68$. Then Algorithm 1 terminates with a $(\frac{(1+5\varepsilon)}{1-\omega}, 1-\delta)$-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 $\widehat{\mu}$: By assumption (A1), we have $|\mathcal{R}_t|_{P_t}| \le g(|P_t|)$, for all t. Thus we can write

$$\Phi(t) = \sum_{P\in\mathcal{R}_t|_{P_t}} (1-\varepsilon)^{|P|} w_0(\mathcal{R}_t[P]), \tag{12}$$

where $\mathcal{R}_t[P] := \{R \in \mathcal{R}_t : \ R \cap P_t = P\}$. Since $\Phi(t)$ satisfies (6), we get by (12) that

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

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

$$\frac{|P_t|}{z^*_{\mathcal{F}}T} \le \frac{(1+\varepsilon)\left(\ln\Phi(0) - \ln(w_0(\mathcal{R}_t[P]))\right)}{\varepsilon(1-\omega)T} + \frac{(1+\varepsilon)|P|}{\varepsilon(1-\omega)T}\cdot\ln\frac{1}{1-\varepsilon}, \qquad \text{for all } P \in \mathcal{R}_t|_{P_t}. \tag{13}$$

Since

$$w_0(\mathcal{R}_t) = w_0\left(\bigcup_{P\in\mathcal{R}_t|_{P_t}} \mathcal{R}_t[P]\right) = \sum_{P\in\mathcal{R}_t|_{P_t}} w_0\left(\mathcal{R}_t[P]\right),$$

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 (13) for $t = t_f - 1$ and $\widehat{P} \in \mathcal{R}_t|_{P_t}$. Using $\Phi(0) = w_0(\mathcal{R}) \le \frac{w_0(\mathcal{R}_t)}{\delta}$, $|\mathcal{R}|_{P_t}| \le g_{\mathcal{F}}(|P_t|) \le g_{\mathcal{F}}(t_{\max})$, $\widehat{\mu}(Q) = \frac{|P_t|+1}{T}$, $|\widehat{P}| < T$ (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

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

for $\varepsilon \le 0.68$. □

5.4 Satisfying the condition on T

As $g_{\mathcal{F}}(t_{\max}) \le t_{\max}^{\gamma}$, for some constant $\gamma \ge 1$ by assumption (A1), and $t_{\max} = \frac{T\cdot\textsc{Opt}_{\mathcal{F}}}{\varepsilon(1-\omega)}\left(\ln\frac{1}{1-\varepsilon} + \ln\frac{1}{\varepsilon\delta}\right)$ as defined in Lemma 13, it is enough to select T to satisfy $T \ge \frac{\gamma\ln T}{\varepsilon^2} + \frac{\gamma\ln T_0}{\varepsilon^2}$, or

$$\frac{T}{a} > \ln T + b, \tag{14}$$

where T_0, a, and b are given by (3).

Set $T = e^C ab(\ln(a+e-1)+1)$, where C is a large enough constant. Then the left-hand side of (14) is $\frac{T}{a} = e^C b\ln(a+e-1) + e^C b$, while the right-hand side is $\ln T + b = C + \ln a + \ln(\ln(a+e-1)+1) + \ln b + b$. Now we need to choose $C \ge 1$ such that $e^C > C + 3$ (say $C = 2$). Then

$$e^C b > 2b > b + \ln b,$$

$$
\begin{aligned}
e^C b\ln(a+e-1) &\ge e^C \log(a+e-1) > (C+3)\ln(a+e-1)\\
&\ge C+1+2\ln(a+e-1) > C+\ln a+\ln(\ln(a+e-1)+1).
\end{aligned}
$$

Thus, setting $T = \Theta(\frac{\gamma}{\varepsilon^2}\log\frac{\gamma}{\varepsilon}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\varepsilon\delta(1-\omega)})$ satisfies the required condition on T.

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 $\omega > 0$ and measure $w : \mathcal{R} \to \mathbb{R}_+$, a point $p \in Q$ such that $\xi_w(p) \ge (1-\omega)\max_{q\in Q}\xi_w(q)$, where $\xi_w(p) := w(\mathcal{R}[p])$.

We will assume here 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 $\mathcal{Q}_{\mathcal{R}'}$. Note by Lemmas 1 and 2 that $|\mathcal{Q}(\mathcal{R}')| \le g(|\mathcal{R}'|, 2^{d+1})$.
- $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$: 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 $\cap_{R\in\mathcal{R}'} R$ (if one exists).
- $\textsc{Sample}(\mathcal{F}, \widehat{w})$: Given $\mathcal{F} = (Q, \mathcal{R})$ and a probability measure $\hat{w} : \mathcal{R} \to \mathbb{R}_+$, it samples from $\widehat{w}$.

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 $\epsilon := \frac{\omega}{2\textsc{Opt}_{\mathcal{F}}}$. By Theorem 5 a random sample $\mathcal{R}'$ of size $N = O(\frac{d2^d}{\epsilon^2}\log\frac{1}{\epsilon}) = O(\frac{d2^d\textsc{Opt}_{\mathcal{F}}^2}{\omega^2}\log\frac{\textsc{Opt}_{\mathcal{F}}}{\omega})$ from $\mathcal{R}$ according to the probability measure $\widehat{w} := w/w(\mathcal{R})$, is an ϵ-approximation with high probability. We call $\textsc{Subsys}(\mathcal{F}^*, \mathcal{R}')$ to obtain the set $Q(\mathcal{R}')$, then return the subset of ranges $\mathcal{R}''' \in \text{argmax}_{\mathcal{R}''\in Q(\mathcal{R}')}|\mathcal{R}''|$. Finally, we call the oracle $\textsc{PointIn}(\mathcal{F}, \mathcal{R}''')$ to obtain a point $p \in \cap_{R\in\mathcal{R}'''} R$.

Lemma 16. *$\xi(p) \ge (1-\omega)\max_{q\in Q}\xi(q)$.*

Proof. The proof, which we include for completeness, goes along the same lines in [12]. Let q^* be a point in $\text{argmax}_{q\in Q}\xi(q)$. Note that by assumption (A2), $w(\mathcal{R}[q^*]) \ge \frac{1}{\textsc{Opt}_{\mathcal{F}}} w(\mathcal{R})$. Then by (1),

$$
\begin{aligned}
\frac{w(\mathcal{R}[p])}{w(\mathcal{R})} &\ge \frac{|\mathcal{R}'[p]|}{|\mathcal{R}'|} - \epsilon \ge \frac{|\mathcal{R}'[q^*]|}{|\mathcal{R}'|} - \epsilon \ge \frac{w(\mathcal{R}[q^*])}{w(\mathcal{R})} - 2\epsilon\\
&= \frac{w(\mathcal{R}[q^*])}{w(\mathcal{R})} - \frac{w}{\textsc{Opt}_{\mathcal{F}}} \ge (1-\omega)\frac{w(\mathcal{R}[q^*])}{w(\mathcal{R})}.
\end{aligned}
$$

The statement follows. □

Remark 1. *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 problems

In the art gallery problem we are given a (non-simple) polygon H with n vertices and h holes, and two sets of points $G, N \subseteq H$. 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 guard all the points in N using candidate guards from G, that is, to find a subset $G' \subseteq G$ such that for every point $q \in N$, there is a point $p \in G'$ such that $p \sim q$.

Let $Q = G$, $\mathcal{R} = \{V_H(q) : \ q \in N\}$, 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 N. We shall see below that the range space $\mathcal{F} = (Q, \mathcal{R})$ satisfies (A1)-(A4).

Related work. Valtr [50] showed that VC-dim$(\mathcal{F}) \leq 23$ for simple polygons and VC-dim$(\mathcal{F}) = O(\log h)$ for polygons with h holes. For simple polygons, this has been improved to VC-dim$(\mathcal{F}) \leq 14$ by Gilbers and Klein [26].

If one of the sets G or N is an (explicitly given) discrete set, then the problem can be easily reduced to a standard SetCover problem. For the case when $G = V$ is the vertex set of the polygon (called *vertex guards*), Ghosh [24, 25] gave an $O(\log n)$-approximation algorithm that runs in time $O(n^4)$ for simple polygons (resp., in time $O(n^5)$ for non-simple polygons). This has been improved by King [33] to $O(\log\log \text{Opt})$-approximation in time $O(n^3)$ for simple polygons (resp., $O((1 + \log(h + 1)) \log \text{Opt})$-approximation in time $O(n^2v^3)$ for non-simple polygons), where Opt here is the size of an optimum set of *vertex* guards. The main ingredient for the improvement in the approximation ratio is the fact proved by King and Kirkpatrick [34] that there is an ϵ-net, in this case and in fact more generally when $G = \partial H$ (called *perimeter guards*), of size $O(\frac{1}{\epsilon} \log\log \frac{1}{\epsilon})$.

For the case when both N and G are infinite, a *discretization* step, which selects a candidate discrete set of guards guaranteed to contain a near optimal-set, seems to be necessary for reducing the problem to SetCover. Such a discretization method was given in [17] that allows an $O(\log \text{Opt})$-approximation for simple polygons (resp., $O(\log h \log(\text{Opt} \cdot \log h))$-approximation in non-simple polygons) in *pseudo-polynomial* time poly(n, Δ), where the *stretch* Δ is defined as the ratio between the longest and shortest distances between vertices of H. However, very recently, an error in one of the claims in [17] was pointed out by Bonnet and Miltzow [6], who also suggested another discretization procedure that results in an $O(\log \text{Opt}_{\mathcal{F}})$-randomized approximation algorithm, after making the following two assumptions

(AG1) vertices of the polygon have integer components, given by their binary representation;

(AG2) no three extensions meet in a point that is not a vertex, where an extension is a line passing through two vertices.

Under these assumptions, it was shown in [6] that one can use a grid Γ of cell size $\frac{1}{D^{O(1)}}$ such that $\text{Opt}_\Gamma = O(\text{Opt}_{\mathcal{F}})$, where Opt_Γ is the size of an optimum set of guards *restricted* to Γ, and D is the diameter of the polygon. Then one can use the algorithm suggested by Efrat and Har-Peled [18] who gave an efficient implementation of the multiplicative weight updates method of Brönnimann and Goodrich [7] in the case when the locations of potential guards are restricted to a dense Γ. More precisely, the authors in [18] gave a *randomized* $O(\log \text{Opt}_\Gamma)$-approximation algorithm for simple polygons (resp., $O(\log h \log(\text{Opt}_\Gamma \cdot \log h))$-approximation in non-simple polygons) in expected time $O(n\text{Opt}_\Gamma^2 \log \text{Opt}_\Gamma \log(n\text{Opt}_\Gamma) \log^2 \Delta)$ (resp., $nh\text{Opt}_\Gamma^3 \,\text{polylog}\, n \log^2 \Delta)$), where Δ here denotes the ratio between the diameter of the polygon and the grid cell size. Note that this would imply a randomized (weakly) polynomial time approximation algorithm for the *unrestricted* guarding case, if one cand show that for a polygon with rational description (of its vertices), there is a near optimal set which has also a rational description. While it is not clear that this is the case in general, the main result in [6] implies that Δ can be chosen, under assumption (AG2), to be $D^{O(1)}$, which implies by (AG1) that $\log \Delta$ is linear in the maximum bit-length of a vertex coordinate. Note that, the same argument combined with Theorem 19 in the appendix shows that one can actually obtain $O(\log z^*_{\mathcal{F}})$-approximation in randomized polynomial-time for simple polygons under assumptions (AG1) and (AG2).

On the hardness side [19], the vertex (and point) guarding problem for simple polygons is known to be APX-hard [19], while the problem for non-simple polygons is as hard as SetCover and hence cannot be approximated by a polynomial time algorithm with ratio less than $((1-\epsilon)/12)\log n)$, for any $\epsilon > 0$, unless $NP \subseteq TIME(n^{O(\log\log n)})$.

7.1.1 Point guards

In this case, we have $Q \leftrightarrow \mathcal{R} \leftrightarrow G = N = H$. Note that (A1) is satisfied with $\gamma = \text{VC-dim}(\mathcal{F}) \leq 14$ by Lemma 1 and the result of [26]. It is also known (see, e.g., [18]) that a subsystem oracle as in (A1$'$) can be computed efficiently, for any (finite) $P \subset Q$, as follows. Let $\mathcal{R}' := \{V_H(p) : p \in P\}$. Then $\mathcal{R}'$ is a finite set of polygons which induces an arrangement of lines (in $\mathbb{R}^2$) of total complexity $O(nh|P|^2)$. We can construct this arrangement in time $O(nh|P|^2 \log(nh|P|))$, and label each cell of the arrangement by the set of visibility polygons it is contained in. Then $\mathcal{R}|_P$ is the set of different cell labels which can be obtained, for e.g., by a sweep algorithm in time $O(nh|P|^2 \log(nh|P|))$.

(A2) follows immediately from the fact that each point in the polygon is seen from some vertex. (A3) is satisfied if we use $w_0 \equiv 1$ to be the area measure over H (recall that ranges in $\mathcal{R}$ are in one-to-one correspondence with points in H). Thus we obtain the following result from Corollary 2, in the *unit-cost* model, since for any $\mathcal{R}' \subseteq \mathcal{R}$, $|Q|_{\mathcal{R}'}| \leq nh|\mathcal{R}|^2$ and hence $g_{\mathcal{F}^*}(r) \leq nhr^2$.

Corollary 3. *Given a polygon H with n vertices and h holes and $\delta > 0$, there is a randomized algorithm that finds in $O(nh^3 \text{OPT}_{\mathcal{F}}^5 \log \frac{\text{OPT}_{\mathcal{F}}}{\delta} \log^2 \text{OPT}_{\mathcal{F}} \log^2(h+2))$ time a set of points in H of size $O(z^*_{\mathcal{F}} \log z^*_{\mathcal{F}} \log(h+2))$ guarding at least $(1-\delta)$ of the area of H, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

We obtain next a deterministic version of Corollary 3 in the *bit model* of computation.

A deterministic maximization oracle. We assume that the components of the vertices have rational representation, with maximum bit-length L for each component (i.e., essentially satisfy (AG1)).

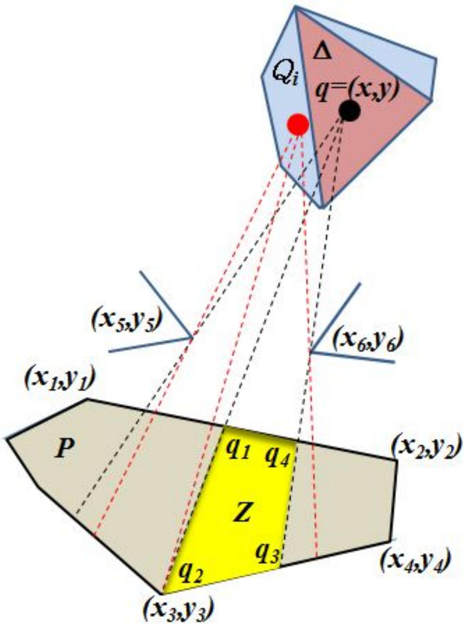
In a given iteration t of Algorithm 1, we are given an active subset of ranges $\mathcal{R}_t \subseteq \mathcal{R}$, determined by the current set of chosen points $P_t \subseteq Q$, and the current measure $w_t : \mathcal{R}_t \to \mathbb{R}_+$, given by $w_t(R) = (1-\varepsilon)^{|P_t \cap R|} w_0(R)$, for $R \in \mathcal{R}_t$, where $w_t(\mathcal{R}_t) \geq \delta \cdot w_0(\mathcal{R})$. Let $\mathcal{R}'_t := \{V_H(p) : p \in P_t\}$. Note that, as explained above, the set of (convex) cells induced by $\mathcal{R}'_t$ over H has complexity (say number of edges) $r_t := O(nh|P_t|^2)$ and can be computed in time[5] $O(nh|P_t|^2 \log(nh|P_t|))$; let us call this set cells$(\mathcal{R}'_t)$, and for any $P \in \text{cells}(\mathcal{R}'_t)$, define $\deg_t(P) := |\{p \in P_t : p \sim q \text{ for some } q \in P\}|$ (recall that all points in P are equivalent w.r.t. visibility from P_t). Note that (the subset of H corresponding to) $\mathcal{R}_t$ can be computed as $\mathcal{R}_t = \bigcup_{P \in \text{cells}(\mathcal{R}'_t) : \deg_t(P) < T} P$. We can write $\xi_t(q) := w_t(\mathcal{R}_t[q])$ for any $q \in Q = H$ as

$$\xi_t(q) = \sum_{P \in \text{cells}(\mathcal{R}'_t)} (1-\varepsilon)^{\deg_t(P)} \text{area}(V_H(q) \cap P). \tag{15}$$

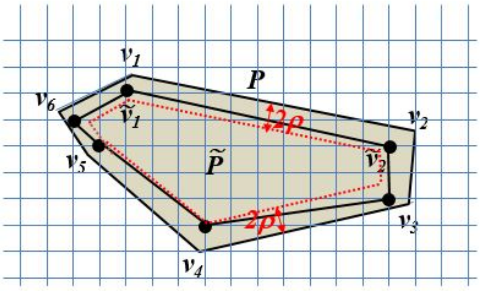
Now, to find the point q in H maximizing $\xi_t(q)$, we follow[6] [43] in expressing $\xi_t(q)$ as a (non-linear) continuous function of two variables, namely, the x and y-coordinates of q. To do this,

[5] For simplicity of presentation, we do not attempt here to optimize the running time, for instance, by maintaining a data structure for computing $\mathcal{R}'_t$, which can be efficiently updated when a new point is added to P_t.

[6] It should be noted that an FPTAS was claimed in [43] when $w_t \equiv 1$, but this claim was not substantiated with a rigorous proof. In fact one of the statements leading to this claim does not seem to be correct, namely that the visibility region of the maximizer q^* can be covered by a *constant* number of points that can be described only in terms of the input description of the polygon.



(a) The visibility region of $q \in \Delta \subseteq Q_i$ inside P can be decomposed into two convex quadrilaterals; one of them is Z, which is determined by $q = (x, y)$, the edges $\{(x_1, y_1), (x_2, y_2)\}$ and $\{(x_3, y_3), (x_4, y_4)\}$ of P, and the vertex (x_6, y_6) of the polygon H.



(b) The polygon P and its internal approximation $\widetilde{P}$ using the vertices of the grid. Note that, since for all vertices v of P the distance between v and $\widetilde{v}$ is at most ρ, the difference $P \setminus \widetilde{P}$ is covered by the region between the ∂P and the dotted red polygon, which is at distance ρ from the boundary of P

we first construct the partition $\{Q_1, \ldots, Q_l\}$ of H, induced by the arrangement of lines formed by the union $\mathcal{V}$ of the vertices of H and the vertices of cells$(\mathcal{R}'_t)$. Note that for any convex cell Q_i in this partition, any two points in Q_i are equivalent w.r.t. the visibility of points from $\mathcal{V}$. Moreover, for any pair of vertices p, p' of a cell $P \in$ cells$(\mathcal{R}'_t)$, any two points in Q_i lie on the same side of the line through p and p'. This implies that, for any point $q = (x, y) \in Q_i \subseteq \mathbb{R}^2$, the set $V_H(q) \cap P$ can be decomposed into at most $|E| = r_t$ regions that are either convex quadrilaterals or triangles, where E is the set of edges of cells$(\mathcal{R}'_t)$; see Figure 1a for an illustration. Using the notation in the figure, we can write the vertices of the quadrilateral Z in counterclockwise order as $q_i = (\frac{a_i(x,y)}{b_i(x,y)}, \frac{c_i(x,y)}{d_i(x,y)})$, for $i = 1, \ldots, 4$, where $a_i(x, y), b_i(x, y), c_i(x, y)$, and $d_i(x, y)$ are affine functions of the form $Ax + By + C$, for some constants $A, B, C \in \mathbb{Q}$ which are multi-linear of degree at most 3 in the components of some of the vertices of P and H. By the *Shoelace formula*, we can further write the area of Z as

$$\text{area}(Z) = \frac{1}{2} \sum_{i=1}^{4} \frac{a_i(x,y)}{b_i(x,y)} \left(\frac{c_{i+1}(x,y)}{d_{i+1}(x,y)} - \frac{c_{i-1}(x,y)}{d_{i-1}(x,y)} \right), \tag{16}$$

where indices wrap-around from 1 to 4. By considering a triangulation of Q_i, and letting Δ be the triangle containing $q \in Q_i$, we can write $q = (x, y) = \lambda_1(x', y') + \lambda_2(x'', y'') + (1 - \lambda_1 - \lambda_2)(x''', y''')$, where $(x', y'), (x', y')$ and (x''', y''') are the vertices of Δ, and $\lambda_1, \lambda_2 \in [0, 1]$. It follows from (15) and (16) that $\xi_t(q)$ can be written as

$$\xi_t(q) = \xi_t(\lambda_1, \lambda_2) = \sum_{i=1}^{k} (1 - \varepsilon)^{j_i} \frac{M_i(\lambda_1, \lambda_2)}{N_i(\lambda_1, \lambda_2)}, \tag{17}$$

where $k = O(|E|) = O(r_t) = \text{poly}(n, h, \log \frac{1}{\delta})$, $j_i \leq |P_t| \leq t_{\max} = \text{poly}(n, h, \log \frac{1}{\delta})$, and $M_i(\lambda_1, \lambda_2)$ and $N_i(\lambda_1, \lambda_2)$ are quadratic functions of λ_1 and λ_2 with coefficients having bit-length $O(L')$, where L' is the maximum bit length needed to represent the components of the vertices in $\mathcal{V}$. We can maximize $\xi_t(\lambda_1,\lambda_2)$ over $\lambda_1,\lambda_2\in[0,1]$ by considering 9 cases, corresponding to $\lambda_1\in\{0,1\}$, and $\lambda_1\in(0,1)$; and $\lambda_2\in\{0,1\}$, and $\lambda_2\in(0,1)$, and taking the value that maximizes $\xi_t(\lambda_1,\lambda_2)$ among them. Consider w.l.o.g. the case when $\lambda_1,\lambda_2\in(0,1)$. We can maximize $\xi_t(\lambda_1,\lambda_2)$ by setting the gradient of (17) to 0, which in turn reduces to solving a system of two polynomial equations of degree $O(k)$ in two variables. A rational approximation to the solution $(\lambda_1^*,\lambda_2^*)$ of this system to within an additive accuracy of τ can be computed in time and bit complexity $\text{poly}(L',k,\log\frac{1}{\tau})$, using, e.g., the *quantifier elimination algorithm* of Renegar [46]; see also Basu et al. [5] and Grigor'ev and Vorobjov [28].

Claim 17. *The function $\xi_t(\lambda_1,\lambda_2)$ in (17) is $2^{O(kL')}$-Lipschitz*[7].

Proof. It is enough to show that $\|\nabla\xi_t(\lambda_1,\lambda_2)\|_2\le 2^{O(kL')}$. By (17), each component of $\nabla\xi_t(\lambda_1,\lambda_2)$ is of the form $\frac{M(\lambda_1,\lambda_2)}{N(\lambda_1,\lambda_2)}$, where $M(\cdot,\cdot)$ and $N(\cdot,\cdot)$ are polynomials in $\lambda_1,\lambda_2\in[0,1]$ of degree $O(k)$ and coefficients of maximum bit length $O(kL')$. Thus $|M(\lambda_1,\lambda_2)|\le 2^{O(kL')}$. Also, from (16) and (17), $N(\lambda_1,\lambda_2)$ can be written as a product of k factors of the form $b(\lambda_1,\lambda_2)^2d(\lambda_1,\lambda_2)^2$, where $b(\cdot,\cdot)$ and $d(\cdot,\cdot)$ can be assumed to be *strictly positive* affine functions of λ_1 and λ_2. Suppose $b(\lambda_1,\lambda_2)=A\lambda_1+B\lambda_2+C$, for some constants $A,B,C\in\mathbb{Q}$ which have bit length $O(L')$. Since the minimum of $b(\lambda_1,\lambda_2)$ over $\lambda_1,\lambda_2\in[0,1]$ is attained at some $\lambda_1,\lambda_2\in\{0,1\}$, it follows that $b(\lambda_1,\lambda_2)\ge\min\{C,A+C,B+C,A+B+C\}\ge\frac{1}{2^{O(L')}}$. A similar observation can be made for $d(\cdot,\cdot)$ and implies that $N(\lambda_1,\lambda_2)\ge\frac{1}{2^{O(kL')}}$, which in turn implies the claim. □

Let $q_\Delta^*\in\text{argmax}_{q\in\Delta}\,\xi_t(q)$. By the above claim, we can choose $\tau=\epsilon 2^{-O(kL')}$ sufficiently small, to get a point $q_\Delta\in\Delta$ such that $\xi_t(q_\Delta)\ge\xi_t(q_\Delta^*)-\epsilon$, where $\epsilon:=\frac{\omega\cdot w_t(\mathcal{R}_t)}{\text{OPT}_{\mathcal{F}}}\ge\frac{\omega\delta(1-\varepsilon)^T w_0(\mathcal{R})}{\text{OPT}_{\mathcal{F}}}$ (and hence $\log\frac{1}{\tau}=\text{poly}(k,L',\log\frac{1}{\delta})$). Finally, we let $p\in\text{argmax}_\Delta\,\xi_t(q_\Delta)$, where Δ ranges over all triangles in the triangulations of $Q_1,\ldots,Q_l$, to get

$$\xi_t(p)\ge\max_{q\in Q}\xi_t(q)-\epsilon=\max_{q\in Q}\xi_t(q)-\frac{\omega\cdot w_t(\mathcal{R}_t)}{\text{OPT}_{\mathcal{F}}}\ge(1-\omega)\max_{q\in Q}\xi_t(q),$$

where the last inequality follows from $\max_{q\in Q}\xi_t(q)\ge\frac{w_t(\mathcal{R}_t)}{\text{OPT}_{\mathcal{F}}}$, implied by (A2).

Rounding. A technical hurdle in the above implementation of the maximization oracle is that the required bit length may grow from one iteration to the next (since the approximate maximizer p above has bit length $\text{poly}(k,L',\log\frac{1}{\tau})$), 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[8]. This can be done as follows. Recall that $\mathcal{R}_t$ can be decomposed by the current set of points P_t into a set $\text{cells}(\mathcal{R}_t)$ of $r_t:=Cnh|P_t|^2$ disjoint convex polygons, for some constant $C>0$. Let $t_{\max}$ be the upper bound on the number of iterations given in Lemma 13, and set $r_{\max}:=Cnht_{\max}^2$. We consider an infinite grid Γ in the plane of cell size $\rho=\frac{\delta\cdot\text{area}(\mathcal{R})}{16Dt_{\max}r_{\max}}$, where D is the diameter of H (which has bit length bounded by $O(L)$).

Let us call a cell $P\in\text{cells}(\mathcal{R}_t)$ *large* if $\text{area}(P)\ge\frac{\delta\cdot\text{area}(\mathcal{R})}{4r_tt_{\max}}$, and *small* otherwise. Let $\mathcal{L}_t$ be the set of large cells in iteration t of the algorithm. For each $P\in\mathcal{L}_t$ we define an approximate polygon $\widetilde{P}\subseteq P$ as follows: for each vertex v of P, we find a point $\widetilde{v}$ in $\Gamma\cap H$, closest to it, then define $\widetilde{P}:=\text{conv. hull}\{\widetilde{v}:\ v\text{ is a vertex of }P\}$. Now, we let $\widetilde{\mathcal{R}}_t:=\bigcup_{P\in\mathcal{L}_t}\widetilde{P}$. The following claim states that the total fraction of ranges that might not be covered due to this approximation is no more than $\delta/2$.

Claim 18. $\sum_{t=1}^{t_f-1}\text{area}(\mathcal{R}_t\setminus\widetilde{\mathcal{R}}_t)\le\frac{\delta}{2}\,\text{area}(\mathcal{R})$.

[7] A continuous differentiable function $f:S\to\mathbb{R}$ is τ-Lipschitz over $S\subseteq\mathbb{R}^n$ if $|f(y)-f(x)|\le\tau\|x-y\|_2$ for all $x,y\in S$.

[8] This is somewhat similar to the rounding step typically applied in numerical analysis to ensure that the intermediate numbers used during the computation have finite precision.

Proof. Two sets contribute to the difference $\mathcal{R}_t \setminus \widetilde{\mathcal{R}}_t$: the set of small cells, and the truncated parts of the larges cells $\bigcup_{P\in\mathcal{L}_t} P \setminus \widetilde{P}$. Note that the total area of the small cells is at most $\sum_{t=1}^{t_f-1} r_t \cdot \frac{\delta\cdot\mathrm{area}(\mathcal{R})}{4r_t t_{\max}} < \frac{\delta}{4}\cdot\mathrm{area}(\mathcal{R})$. On the other hand, for any $P \in \mathcal{L}_t$, we have $\mathrm{area}(P) - \mathrm{area}(\widetilde{P}) \le 2\rho \cdot \mathrm{prem}(P)$, where $\mathrm{prem}(P)$ is the length of the perimeter of P. This inequality holds because $P \setminus \widetilde{P}$ is contained in the region at distance 2ρ from the boundary of P; see Figure 1b for an illustration. It follows that

$$\sum_{t=1}^{t_f-1} \sum_{P\in\mathcal{L}_t} \mathrm{area}(P \setminus \widetilde{P}) \le 2\rho \cdot \sum_{t=1}^{t_f-1} \sum_{P\in\mathcal{L}_t} \mathrm{prem}(P) \le 4\rho \cdot \sum_{t=1}^{t_f-1} r_t D < 4\rho t_f r_{t_f} D \le \frac{\delta}{4} \cdot \mathrm{area}(\mathcal{R}),$$

by our selection of ρ. The claim follows. □

The only change we need in Algorithm 1 is to replace $\mathcal{R}_t$ in by $\widetilde{\mathcal{R}}_t$. (It is easy to see that the analysis also goes through with almost no change; we just have to replace $\mathcal{R}_t$ by $\widetilde{\mathcal{R}}_t$ and δ by $\frac{\delta}{2}$.)

Note now that, since the polygon is contained in a square of size $2D$, the total number of points in Γ we need to consider is at most

$$\frac{2D}{\rho} = \frac{32D^2 t_{\max} r_{\max}}{\delta \cdot \mathrm{area}(H)} = 2^{O(L)}\,\mathrm{poly}(n, h, \frac{1}{\delta}),$$

and thus the number of bits needed to represent each point of Γ is $L \cdot \mathrm{polylog}(n, h, \frac{1}{\delta})$. Since the vertices of each cell $\widetilde{P}$ lie on the grid, the bit length L' used in the computations above (in the implementation of the maximization oracle) and the overall running time is $\mathrm{poly}(L, n, h, \log\frac{1}{\delta})$.

Corollary 4. *Given a simple polygon H with n vertices with rational representation of maximum bit-length L and $\delta > 0$, there is a deterministic algorithm that finds in* $\mathrm{poly}(L, n, \log\frac{1}{\delta})$ *time a set of points in H of size $O(z^*_{\mathcal{F}} \log z^*_{\mathcal{F}} \log(h+2))$ and bit complexity* $\mathrm{poly}(L, n, \log\frac{1}{\delta})$ *guarding at least $(1-\delta)$ of the area of H, where $z^*_{\mathcal{F}}$ is the value of the optimal fractional solution.*

If H is not simple, we get a result similar to Corollary 4 but with a quasi-polynomial running time $\mathrm{poly}(L, n^{O(\log h)}, \log\frac{1}{\delta})$ (due to the complexity of the deterministic ϵ-net finder).

Remark 2. *It is worth noting that one can also obtain a* randomized *approximation algorithm with the same guarantee of Corollary 5 from the results in [6], by first randomly perturbing the polygon H into a new polygon H' such that $H' \subseteq H$ and* $\mathrm{area}(H \setminus H') \le \delta$. *Such a perturbation can be done using the rounding idea described above and guarantees with high probability that (AG2) is satisfied. Thus, we can apply the result in [6] on H'.*

7.1.2 Perimeter guards

In this case, we have $Q \leftrightarrow G = \partial H$ and $\mathcal{R} \leftrightarrow T = H$. This is similar to the point guarding case with the exception that, in the maximization oracle, the point q in (17) is selected from a line segment on ∂H. Also, by [34], the range space in this case admits an ϵ-net of size $O(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon})$. Thus we get the following result.

Corollary 5. *Given a simple polygon H with n vertices with rational representation of maximum bit-length L and $\delta > 0$, there is a deterministic algorithm that finds in* $\mathrm{poly}(L, n, h, \log\frac{1}{\delta})$ *time a set of points in ∂H of size $(z^*_{\mathcal{F}} \log\log z^*_{\mathcal{F}})$ and bit complexity* $\mathrm{poly}(L, n, h, \log\frac{1}{\delta})$ *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 [30] and generalizations in 3D [16, 36]. Fewer results are known for the continuous case, e.g., [27] which considers the covering of simple polygons by translates of a rectangle[9] and only provides an exact (exponential-time) algorithm; see also [22] 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 $\mathcal{R} := \big\{\{H_0^i \in Q :\ R \in H_0^i\} :\ R \in \bigcup_{H\in\mathcal{P}} H\big\}$. 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 [44] that VC-dim$(\mathcal{F}^*) \leq 3$ and also that $\mathcal{F}^*$ admits an ϵ-net of size $s_{\mathcal{F}^*} = O(\frac{1}{\epsilon})$. As observed in [36], this would also imply that VC-dim$(\mathcal{F}) \leq 3$ and $s_{\mathcal{F}} = O(\frac{1}{\epsilon})$. Thus (A1) is satisfied with $\gamma = 3$; also we can show that (A1$'$) is satisfied as follows. Let m be the total number of vertices of the polygons in $\mathcal{H}$ and H_0. Given a finite subset $P \subseteq Q$ 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 $\mathcal{H} \cup P$ (where a cell is naturally defined to be a maximal set of points in $\mathcal{R}$ that all belong exactly to the same polygons in the arrangement). Let us cal this set cells$(\mathcal{R})$ and note that it has size $O(m)$. Note also that every cell $\mathcal{R}' \in$ cells$(\mathcal{R})$ is labeled by the subset $S(\mathcal{R}')$ of P that contains it, and $\mathcal{R}|_P$ is the set of different labels.

Assume 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 $\textsc{Opt}_{\mathcal{F}} \leq \frac{D}{d}$. (A3) is satisfied if we use $w_0 \equiv 1$ to be 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. Note that for a given finite $\mathcal{R}' \subseteq \mathcal{R}$, the set $Q_{\mathcal{R}'}$ 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 $O((m|\mathcal{R}'|)^2)$ time. This argument also shows that $\textsc{PointIn}(\mathcal{F}, \mathcal{R}')$ can be implemented in the time $\emptyset((m|\mathcal{R}'|)^2)$. Finally, we can implement $\textsc{Sample}(\mathcal{F}, \widehat{w}_t)$ given the probability measure $\widehat{w}_t : \mathcal{R} \to \mathbb{R}_+$ defined by the subset $P_t \subseteq Q$ as follows. We construct the cell arrangement cells$(\mathcal{R})$, induced by $P = P_t$ as described above. We first sample $\mathcal{R}'$ with probability $\frac{\widehat{w}_t(\mathcal{R}')}{\sum_{\mathcal{R}'\in \text{cells}(\mathcal{R})} \widehat{w}_t(\mathcal{R}')w_0(\mathcal{R}')}$, then we sample a point R uniformly at random from $\mathcal{R}'$.

Corollary 6. *Given a collection of polygons in the plane $\mathcal{H}$ be and a (full-dimensional) convex polygon H_0, with m total vertices respectively and $\delta > 0$, there is a randomized algorithm that finds in* 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 a hitting set problem in a range space $\mathcal{F} = (Q, \mathcal{R})$, where Q is the set of supporting hyperplanes for P_1 and $\mathcal{R} := \{\{p \in Q :\ p \text{ separtaes } R \text{ from } \mathcal{P}_1\} :\ R \in \partial\mathcal{P}_2\}$. Note that VC-dim$(\mathcal{F}) = d$ (and VC-dim$(\mathcal{F}^*) = d+1$). In their paper [7], Brönnimann and

[9] Note that in [27], each polygon has to be covered *completely* by a rectangle.

Goodrich gave a deterministic $O(d^2 \log \textsc{Opt}_{\mathcal{F}})$-approximation algorithm, improving on earlier results by Mitchell and Suri [42], and Clarkson [15]. It was shown in [42] 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. Clearly (A1) is satisfied with $\gamma = d$, and given a finite set of hyperplanes $P \subseteq Q$ we can find the projection $\mathcal{R}|_P$ as follows. We first construct the cells of the hyperplane arrangement of P, which has complexity $O(|P|^d)$, in time $O(|P|^{d+1})$; see, e.g., [4, 49]. 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 $\text{cells}(\mathcal{R})$ (recall that $\mathcal{R} \leftrightarrow \partial\mathcal{P}_2$). Every $\mathcal{R}' \in \text{cells}(\mathcal{R})$ can be identified with the subset $S(\mathcal{R}')$ of P that separates a point $R \in \mathcal{R}'$ from $\mathcal{P}_1$. Then $\mathcal{R}|_P = \{S(\mathcal{R}') : \mathcal{R}' \in \text{cells}(\mathcal{R})\}$. The running time for this is $\text{poly}(|P|^d, m^d)$. Also, (A2) is obviously satisfied since $\mathcal{P}_3 = \mathcal{P}_2$ is a separator with n facets. For (A3), we use the $w_0 \equiv 1$ to be 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 $Q_{\mathcal{R}'}$ 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 $Q_{\mathcal{R}'}$ 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 measure $\widehat{w}_t : \mathcal{R} \to \mathbb{R}_+$ defined by the subset $P_t \subseteq Q$ as follows. We construct the cell arrangement $\text{cells}(\mathcal{R})$, induced by $P = P_t$ as described above. We first sample $\mathcal{R}'$ with probability $\frac{\widehat{w}_t(\mathcal{R}')}{\sum_{\mathcal{R}' \in \text{cells}(\mathcal{R})} \widehat{w}_t(\mathcal{R}') w_0(\mathcal{R}')}$, 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 7. *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 $O(z^*_{\mathcal{F}} \cdot d \log z^*_{\mathcal{F}})$ 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.*

Note that the results in corollaries 6 and 7 assume the unit-cost model of computation and infinite precision arithmetic. We believe that deterministic algorithms for the maximization oracle in the bit-model can also be obtained using similar techniques as in Section 7.1. We leave the details for the interested reader.

Acknowledgement. The author is grateful to Waleed Najy for his help in the proof of Lemmas 15 and 16 and for many useful discussions.

References

[1] Pankaj K. Agarwal and Jiangwei Pan. Near-linear algorithms for geometric hitting sets and set covers. In *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, pages 271:271–271:279, New York, NY, USA, 2014. ACM.

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

[3] Boris Aronov, Esther Ezra, and Micha Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM Journal on Computing*, 39(7):3248–3282, 2010.

[4] David Avis and Komei 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] Edouard Bonnet and Tillmann Miltzow. An approximation algorithm for the art gallery problem. In *uropean Workshop on Computational Geometry*, EuroCG'16, 2016.

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

[8] Hervé Brönnimann, Bernard Chazelle, and Jiři Matoušek. Product range spaces, sensitive sampling, and derandomization. *SIAM Journal on Computing*, 28(5):1552–1575, 1999.

[9] Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 1576–1585. SIAM, 2012.

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

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

[12] Otfried Cheong, Alon Efrat, and Sariel Har-Peled. On finding a guard that sees most and a shop that sells most. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '04, pages 1098–1107, Philadelphia, PA, USA, 2004. Society for Industrial and Applied Mathematics.

[13] Otfried Cheong, Alon Efrat, and Sariel Har-Peled. Finding a guard that sees most and a shop that sells most. *Discrete & Computational Geometry*, 37(4):545–563, 2007.

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

[15] Kenneth L. Clarkson. *Algorithms for polytope covering and approximation*, pages 246–252. Springer Berlin Heidelberg, Berlin, Heidelberg, 1993.

[16] L. Kenneth Clarkson and Kasturi Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete & Computational Geometry*, 37(1):43–58, 2006.

[17] Ajay Deshpande, Taejung Kim, Erik D. Demaine, and Sanjay E. Sarma. *Algorithms and Data Structures: 10th International Workshop, WADS 2007, Halifax, Canada, August 15-17, 2007. Proceedings*, chapter A Pseudopolynomial Time O(logn)-Approximation Algorithm for Art Gallery Problems, pages 163–174. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

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

[19] S. Eidenbenz, C. Stamm, and P. Widmayer. Inapproximability results for guarding polygons and terrains. *Algorithmica*, 31(1):79–113, 2001.

[20] Guy Even, Dror Rawitz, and Shimon (Moni) Shahar. Hitting sets when the vc-dimension is small. *Inf. Process. Lett.*, 95(2):358–362, July 2005.

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

[22] Shashidhara Krishnamurthy Ganjugunte. *Geometric Hitting Sets and Their Variants.* PhD thesis, Duke University, USA, 2011.

[23] N. Garg and J. Könemann. Faster and simpler algorithms for multicommodity flow and other fractional packing problems. In *Proc. 39th Symp. Foundations of Computer Science (FOCS)*, pages 300–309, 1998.

[24] Subir Kumar Ghosh. Approximation algorithms for art gallery problems in polygons. *Canad. Information Processing Soc. Congress.*, pages 429–434, 1987.

[25] Subir Kumar Ghosh. Approximation algorithms for art gallery problems in polygons. *Discrete Applied Mathematics*, 158(6):718 – 722, 2010.

[26] Alexander Gilbers and Rolf Klein. A new upper bound for the vc-dimension of visibility regions. *Computational Geometry*, 47(1):61 – 74, 2014.

[27] Roland Glück. Covering polygons with rectangles. In *uropean Workshop on Computational Geometry*, EuroCG'16, 2016.

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

[29] David Haussler and Emo Welzl. epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.

[30] Dorit S. Hochbaum and Wolfgang Maass. Approximation schemes for covering and packing problems in image processing and vlsi. *J. ACM*, 32(1):130–136, January 1985.

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

[32] A. Kalai and S. Vempala. Simulated annealing for convex optimization. *Mathematics of Operations Research*, 31(2):253–266, 2006.

[33] James King. Fast vertex guarding for polygons with and without holes. *Comput. Geom.*, 46(3):219–231, 2013.

[34] James King and David Kirkpatrick. Improved approximation for guarding simple galleries from the perimeter. *Discrete & Computational Geometry*, 46(2):252–269, 2011.

[35] János Komlós, János Pach, and Gerhard Woeginger. Almost tight bounds for &egr;-nets. *Discrete Comput. Geom.*, 7(2):163–173, March 1992.

[36] Sören Laue. Geometric set cover and hitting sets for polytopes in r^3. In *STACS 2008, 25th Annual Symposium on Theoretical Aspects of Computer Science, Bordeaux, France, February 21-23, 2008, Proceedings*, pages 479–490, 2008.

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

[38] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *J. ACM*, 41(5):960–981, September 1994.

[39] Jiři Matoušek. Cutting hyperplane arrangements. *Discrete & Computational Geometry*, 6(3):385–406, 1991.

[40] Jiři Matoušek. Reporting points in halfspaces. *Computational Geometry*, 2(3):169 – 186, 1992.

[41] Jiří Matoušek, Raimund Seidel, and E. Welzl. How to net a lot with little: Small &egr;-nets for disks and halfspaces. In *Proceedings of the Sixth Annual Symposium on Computational Geometry*, SCG '90, pages 16–22, New York, NY, USA, 1990. ACM.

[42] Joseph S.B. Mitchell and Subhash Suri. Separation and approximation of polyhedral objects. *Computational Geometry*, 5(2):95 – 114, 1995.

[43] Simeon Ntafos and Markos Tsoukalas. Optimum placement of guards. *Information Sciences*, 76(12):141 – 150, 1994.

[44] János Pach and Gerhard Woeginger. Some new bounds for epsilon-nets. In *Proceedings of the Sixth Annual Symposium on Computational Geometry*, SCG '90, pages 10–15, New York, NY, USA, 1990. ACM.

[45] Evangelia Pyrga and Saurabh Ray. New existence proofs ϵ-nets. In *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, SCG '08, pages 199–207, New York, NY, USA, 2008. ACM.

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

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

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

[49] Nora H. Sleumer. Output-sensitive cell enumeration in hyperplane arrangements. *Nordic J. of Computing*, 6(2):137–147, June 1999.

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

[51] 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.

[52] Kasturi Varadarajan. Epsilon nets and union complexity. In *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*, SCG '09, pages 11–16, New York, NY, USA, 2009. ACM.

A An extension of the Brönnimann-Goodrich algorithm for continuous range spaces

In addition to (A1), we will make the following assumption in this section:

(A3′) There exists a finite measure $\mu_0 : P \to \mathbb{R}_+$ such that the ranges in $\mathcal{R}$ are μ_0-measurable.

```
   Data: A range space $\mathcal{F} = (Q, \mathcal{R})$ satsfying (A1) and (A3′), and an approximation
         accuracy $\epsilon \in (0, 1)$.
   Result: A hitting set for $\mathcal{R}$.
 1 $t \leftarrow 0$
 2 repeat
 3     define the probability measure $\widehat{\mu}_t : Q \to \mathbb{R}_+$ by $\widehat{\mu}_t(q) \leftarrow \frac{\mu_t(q)}{\mu_t(Q)}$, for $q \in Q$
 4     Find an $\epsilon$-net $P$ for $\mathcal{R}$ w.r.t. the probability measure $\widehat{\mu}_t$
 5     if there is a range $R_{t+1} \in \mathcal{R}$ such that $R_{t+1} \cap P = \emptyset$ then
 6         $\mu_{t+1}(q) := 2\mu_t(q)$ for all $q \in R_{t+1}$
 7     end
 8     $t \leftarrow t + 1$
 9 until P is a hitting set for $\mathcal{R}$;
10 return $P$
```

Algorithm 2: The Brönnimann-Goodrich hitting set algorithm

For $\mathcal{R}' := \{R_1, \ldots, R_t\} \subseteq \mathcal{R}$, let $\text{cells}(\mathcal{R}) := \{\bigcap_{t' \in S} R_{t'} \setminus \bigcap_{t' \in [t] \setminus S} R_{t'} : \; S \subseteq [t]\} \setminus \{\emptyset\}$ be the partition of Q induced by $\mathcal{R}'$. Define

$$\delta_0 := \min_{\text{finite } \mathcal{R}' \subseteq \mathcal{R}} \; \min_{P \in \text{cells}(\mathcal{R}')} \frac{\mu_0(P)}{\mu_0(Q)}. \tag{18}$$

Theorem 19. *Let $\mathcal{F} = (Q, \mathcal{R})$ be a range space satisfying (A1) and (A3′) and admitting an ϵ-net of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$, and $\mu : Q \to \mathbb{R}_+$ be a measure feasible for (*F-HITTING*). For $\epsilon \leq \frac{1}{2\mu(Q)}$, Algorithm 2 finds a hitting set of size $s_{\mathcal{F}}(\frac{1}{\epsilon})$ in $O(\mu(Q) \log \frac{1}{\delta_0})$ iterations.*

Proof. Let $\mathcal{R}_t = \{R_1, \ldots, R_t\} \subseteq \mathcal{R}$ be the set of ranges whose weights are doubled in iterations $1, \ldots, t$. For $P \subseteq Q$, define $\deg_t(P) := |\{R \in \mathcal{R}_t : \; R \supseteq P\}|$. For two measures $\mu', \mu'' : Q \to \mathbb{R}_+$, denote by $\langle \mu', \mu'' \rangle$ the inner product: $\langle \mu', \mu'' \rangle := \int_{q \in Q} \mu'(q) \mu''(q) dq$. Then

$$\langle \mu_t, \mu \rangle = \sum_{P \in \text{cells}(\mathcal{R}_t)} 2^{\deg_t(P)} \mu(P). \tag{19}$$

By the feasibility of μ, for every $R_{t'} \in \mathcal{R}_t$, we have that $\mu(R_{t'}) = \int_{q \in Q} \mu(q) \mathbf{1}_{q \in R_{t'}} dq \geq 1$. Thus,

$$\begin{aligned} t \leq \sum_{t'=1}^{t} \mu(R_{t'}) &= \sum_{t'=1}^{t} \int_{q \in Q} \mu(q) \mathbf{1}_{q \in R_{t'}} dq \\ &= \int_{q \in Q} \sum_{t'=1}^{t} \mu(q) \mathbf{1}_{q \in R_{t'}} dq = \int_{q \in Q} \mu(q) \deg_t(q) dq \\ &= \sum_{P \in \text{cells}(\mathcal{R}_t)} \deg_t(P) \mu(P). \end{aligned} \tag{20}$$

From (19) and (20), we obtain

$$\frac{\langle \mu_t, \mu \rangle}{\mu(Q)} = \sum_{P \in \text{cells}(\mathcal{R}_t)} 2^{\deg_t(P)} \frac{\mu(P)}{\mu(Q)} \geq 2^{\sum_{P \in \text{cells}(\mathcal{R}_t)} \deg_t(P) \frac{\mu(P)}{\mu(Q)}} \geq 2^{t/\mu(Q)}, \tag{21}$$

where the first inequality follows by the convexity of the exponential function while the second follows from (20). Since the range R_{t+1} chosen in step 5 does not intersect the ϵ-net P chosen in step 4, we have $\mu_t(R_{t+1}) < \epsilon \mu_t(Q)$ and thus $\mu_{t+1}(Q) = \mu_t(Q) + \mu_t(R_{t+1}) < (1 + \epsilon) \mu_t(Q)$. It follows that

$$(1 + \epsilon)^t > \frac{\mu_t(Q)}{\mu_0(Q)} = \sum_{P \in \text{cells}(\mathcal{R}_t)} 2^{\deg_t(P)} \frac{\mu_0(P)}{\mu_0(Q)}. \tag{22}$$

From (21), we get that there is a $P \in \text{cells}(\mathcal{R}_t)$ such that $2^{\deg_t(P)} \geq 2^{t/\mu(Q)}$. On the other hand, (22) implies that $2^{\deg_t(P)}\frac{\mu_0(P)}{\mu_0(Q)} < (1+\epsilon)^t < e^{\epsilon t}$. Putting the two inequalities together, we obtain

$$\frac{t\ln 2}{\mu(Q)} \leq \epsilon t + \ln\frac{\mu_0(Q)}{\mu_0(P)}. \tag{23}$$

Since $\epsilon \leq \frac{1}{2\mu(Q)}$, we get from (23) that $t \leq \frac{1}{\ln 2 - 0.5}\cdot \mu(Q)\ln\frac{\mu_0(Q)}{\mu_0(P)}$. □

Let μ^* be a $(1+\varepsilon)$-approximate solution for (F-HITTING). We can use Algorithm 2 in a binary search manner to determine whether or not $\mu^*(Q) \leq (1+\rho)^i$, for any $\rho > 0$ and $i \in \mathbb{Z}_+$, by checking if the algorithm stops with a hitting set in $\frac{1}{\ln 2-0.5}\cdot(1+\rho)^i\log\frac{1}{\delta_0}$ iterations. As $1 \leq \mu^*(Q) \leq n$ if we assume (A2), we need only $O(\log_{1+\rho} n)$ binary search steps.

We mention an application of Theorem 19 when Q is finite. Let $\mu_0 \equiv 1$. Then $\delta_0 \geq \frac{1}{|Q|}$, and Theorem 19 implies that a hitting set of size $s_{\mathcal{F}}(O(\mu^*(Q)))$ can be found in $O(\mu^*(Q)\log n \log|Q|)$ iterations.

Remark 3. *One can also extend the second algorithm and analysis suggested in [1] to the infinite case, to get as randomized algorithm that computes, with probability at least $\frac{1}{7}$ a hitting set of size $s_{\mathcal{F}}(8z^*_{\mathcal{F}})$ in $O(z^*_{\mathcal{F}}\ln(\frac{1}{\delta_0(\delta'_0)^2}))$, where δ_0 is as defined in (18), and*

$$\delta'_0 := \min_{\textit{finite } P\subseteq Q}\frac{w_0(\mathcal{R}[P])}{w_0(\mathcal{R})}, \tag{24}$$

where $\mathcal{R}[P] := \{R \in \mathcal{R} :\ R\cap Q = P\}$.

enddocument

SODA 2017 reviews

——————— REVIEW 1 ———————

PAPER: 420 TITLE: Finding Small Hitting Sets in Infinite Range Spaces of Bounded VC-dimension AUTHORS: Khaled Elbassioni

———— OVERALL EVALUATION ————

Summary: A decent paper, but the writing is bad and the result is somewhat too weak for SODA.

What is in the paper:

This paper present a way to get a better approximation for partial set-cover for set systems with bounded VC dimension. Formally, if a set system has VC dimension d, and optimal cover with k sets, it is known that one can compute a cover of size $O(dklogk)$, by solving LP and using the eps-net theorem. The author shows how one can cover all by $(1-\delta)$ fraction of the ground set, using $O(dklogk)$ sets.

By itself this does not look too interesting, but when dealing with infinite range spaces this is somewhat more interesting. The main such example is the (general) art-gallery problem, where given a polygon one wants to pick guards that sees the whole polygon. There is some previous work on the approximation version (which is still, BTW, widely open). In particular, a previous work showed (implicitly) that using a grid one has the phenomena that the measure of points that are not guarded appropriately decreases quickly.

What the author does is to revisit the reweighting algorithm for this problem and shows that the measure of points that are not seen sufficient number of times is tiny (smaller than delta). This is a nice observation, and importantly the number of iterations the algorithm does is only polynomial in $\log(1/\delta)$. This is a cute argument (which is partially new) that can be done in half a page. Once you have that, the approximation follows immediately by computing an appropriate net with this new measure of the ground set.

Why I am not excited

My feeling is that nobody really worked on this problem (cover all but delta fraction), and the main insights are present (implicitly and in weakened form) in previous work, which appeared in weaker venues than SODA.

Furthermore, there are several ways to do what the author do, which the author should compare to:

0. Using standard approximation results, implies that one can solve the problem on the set system induced on samples of size $O(d/\delta^2)$.

1. Using relative approximations (http://arxiv.org/abs/0909.0717), computing a sample of size $O((dopt/(\varepsilon^2\delta))\log(1/\delta))$ from the range space, and computing an optimal cover for it, yields the desired approximation. The running time is inferior to what the authors offers (as far as dependency on delta - it is polynomial in 1/delta instead of in the $\log(1/\delta)$) [the only trick here is that the sample is for the complement range that is not being covered].

This yields the same approximation to the author and the description is easy. The author needs to include this in the paper.

BTW, by doing this sample implicitly and running the algorithm on this implicit sample, one can get one the author does - the details are not trivial, and arguably not as interesting as what the author does.

2. Just doing greedy coverage, implies that after $O(k\log(1/\delta))$ iterations all but delta fraction of the range space would be covered. The running time is better than the authors algorithm, but the approximation is potentially worse (i.e., $O(klog(1/\delta)))$ instead of $O(dk\log k)$ [d is the associated VC dimension].

3. The writing is bad - the author over formalizes what he does for no real reason (integrals? really). The basic argument is interesting, and the author should present it succinctly and clearly - it is enough to show the argument for a finite range space and state that it extends readily to the continuous case.

—————— REVIEW 2 ——————

PAPER: 420 TITLE: Finding Small Hitting Sets in Infinite Range Spaces of Bounded VC-dimension AUTHORS: Khaled Elbassioni

——— OVERALL EVALUATION ———

This paper considers the following abstract problem: given a set system $(X, \mathcal{R})$ where the base set of elements X are infinitely many, then compute a small subset of X that hits nearly all the (infinitely many) subsets of $\mathcal{R}$ (more precisely, at least $(1-\delta)*|\mathcal{R}|$ subsets for an input parameter delta).

Specific instances of this problem that are presented in the paper are the following algorithmic results: i) art gallery problem, where the previous best bound of a log $(1/\delta)$ approximation are improved to a log (opt) approximation, ii) covering nearly all of a polygonal region by translates of objects, and iii) separating nearly all of an inner polytope from a containing outer polytope.

The main addition to earlier techniques is that the set system is not finite, to which the earlier ideas do not extend. On the other hand, the previous work (BG, LP rounding) compute a hitting set for all of $\mathcal{R}$, while here a constant fraction of $\mathcal{R}$ is left unhit.

The main idea of the paper is to extend basically the framework of Agarwal-Pan and Grigoriadis-Kachiyan to this case. In the tradition of these weighing methods, it is an iterative algorithm that maintains weights on the sets, and then adds a point hitting the maximum weight sets to a collection in each iteration. At the end the collection of points gives a measure on the points which is then rounded using epsilon-nets.

In general the paper is written very well. At the few places the writing can be a little more precise (in several lemmas the input parameters are not defined, ” should be “, and so on).

The ideas in this paper build nicely on earlier work; actually I am surprised that this work was not done earlier, as the infinite case is a fundamental problem of importance, and various specific problems that this paper improves have been studied reasonably well earlier. For this reason, I recommend acceptance.

——————— REVIEW 3 ———————

PAPER: 420 TITLE: Finding Small Hitting Sets in Infinite Range Spaces of Bounded VC-dimension AUTHORS: Khaled Elbassioni

———— OVERALL EVALUATION ————

Given (a possibly infinite) ground set Q, and a collection of subsets cR ("ranges") of these. A set H of points in Q is a hitting set if H contains a point from each of the given ranges. The authors consider range spaces $(Q, \mathcal{R})$ of bounded VC-dimension d. For such range spaces on finite ground sets Q it was known that one can find $O(d\log(opt))$- and even $O(d\log(opt^*))$-approximate hitting sets (see [7] and [20], resp), where opt and opt^* are integral and fractional optimum hitting sets.

This paper considers range spaces with infinite ground sets. Under certain assumptions, for $\delta > 0$, the authors present an algorithm that finds a subset of points of Q of size $O(dopt^* \log(opt^*))$ that hits a $(1 - \delta)$ fractions of the ranges in $\mathcal{R}$. The algorithm runs in time polynomial in the input description of the range space, and $\log(1/\delta)$.

The authors first formulate a natural convex program for (a more general) version of the problem described above. It is then shown how this convex program can be solved via the multiplicative weights method, assuming that a certain type of oracle exists that is used to pick out a point from Q to add to the partial cover in each iteration.

The last portion of the paper applies the algorithm produced to a number of problems: (a) the art gallery problem, (b) the problem of covering a polygonal region by translates of a convex polygon, and (c) polyhedral separation in fixed dimension. In each case, the authors improve over previous results for the respective problems that made structural assumptions, and hence considered special cases of the problem.

I found the results presented in this paper interesting, but am wondering how big a step forward it is from existing work in the area. In particular, the authors are of course not the first to apply multiplicative weights to covering problems (and also those that have infinte-sized, but oracle-given variable spaces). E.g., Garg & Koenemann had previously applied the method to such problems (e.g., see the Master's thesis of Koenemann).

Also, I am wondering how restrictive the assumptions are that the authors are making. E.g., assumption A1 seems to be a strong one. The assumption says that the projection of the given range set to any point set of size at most r has size bounded by r^γ for some universal constant γ. Does this not imply, e.g., that the given range space has bounded shallow cell complexity?