Small-size ε-Nets for Axis-Parallel Rectangles and Boxes*

Boris Aronov[†] Esther Ezra[‡] Micha Sharir[§]

March 28, 2009

Abstract

We show the existence of ε-nets of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$ for planar point sets and axis-parallel rectangular ranges. The same bound holds for points in the plane and "fat" triangular ranges, and for point sets in $\mathbb{R}^3$ and axis-parallel boxes; these are the first known non-trivial bounds for these range spaces. Our technique also yields improved bounds on the size of ε-nets in the more general context considered by Clarkson and Varadarajan. For example, we show the existence of ε-nets of size $O\left(\frac{1}{\varepsilon}\log\log\log\frac{1}{\varepsilon}\right)$ for the dual range space of "fat" regions and planar point sets (where the regions are the ground objects and the ranges are subsets stabbed by points). Plugging our bounds into the technique of Brönnimann and Goodrich, we obtain improved approximation factors (computable in randomized polynomial time) for the HITTING SET or the SET COVER problems associated with the corresponding range spaces.

*Work on this paper by Boris Aronov and Micha Sharir has been supported by a joint grant from the U.S.-Israel Binational Science Foundation. Work by Boris Aronov has also been supported by NSA MSP Grant H98230-06-1-0016 and NSF Grant CCF-08-30691. Work by Esther Ezra has been supported by ARO grants W911NF-04-1-0278 and W911NF-07-1-0376, and by an NIH grant 1P50-GM-08183-01. Work by Micha Sharir has also been supported by NSF Grants CCF-05-14079 and CCF-08-30272, by Grant 155/05 from the Israel Science Fund, and by the Hermann Minkowski–MINERVA Center for Geometry at Tel Aviv University.

[†]Department of Computer Science and Engineering, Polytechnic Institute of NYU, Brooklyn, NY 11201-3840, USA; *aronov@poly.edu*.

[‡]Department of Computer Science, Duke University, Durham, NC 27708-0129, USA; *esther@cs.duke.edu*.

[§]School of Computer Science, Tel Aviv University, Tel Aviv 69978 Israel and Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA; *michas@post.tau.ac.il*.

1 Introduction

Since their introduction in 1987 by Haussler and Welzl [HW87] (see also Clarkson [Cla87] and Clarkson and Shor [CS89] for related techniques), ε-nets have become one of the central concepts in computational and combinatorial geometry, and have been used in a variety of applications, such as range searching, geometric partitions, and bounds on curve-point incidences, to name a few; see, e.g., Matoušek [Mat02]. We recall their definition: A *range space* $(X, \mathcal{R})$ is a pair consisting of an underlying universe X of objects, and a certain collection $\mathcal{R} \subseteq 2^X$ of subsets (*ranges*). Of particular interest are range spaces of *finite VC-dimension*; the reader is referred to [HW87] for the exact definition. Informally, it suffices to require that, for any finite subset $P \subset X$, the number of distinct sets $r \cap P$, for $r \in \mathcal{R}$, be $O(|P|^d)$, for some constant d (which is upper bounded by the VC-dimension of $(X, \mathcal{R})$).

Given a range space $(X, \mathcal{R})$, a finite subset $P \subset X$, and a parameter $0 < \varepsilon < 1$, an *ε-net for P and $\mathcal{R}$* is a subset $N \subseteq P$ with the property that any range $r \in \mathcal{R}$ with $|r \cap P| \geq \varepsilon |P|$ contains an element of N. In other words, N is a hitting set for all the "heavy" ranges.

The epsilon-net theorem of Haussler and Welzl asserts that, for any $(X, \mathcal{R})$, P, and ε as above, such that $(X, \mathcal{R})$ has finite VC-dimension d, there exists an ε-net N of size $O\left(\frac{d}{\varepsilon} \log \frac{d}{\varepsilon}\right)$, and that in fact a random sample of P of that size is an ε-net with constant probability. In particular, the size of N is independent of the size of P. The bound on the size of the ε-net was later improved to $O\left(\frac{d}{\varepsilon} \log \frac{1}{\varepsilon}\right)$ by Blumer *et al.* [BEHW89], and then to $(1 + o(1))\frac{d}{\varepsilon} \log \frac{1}{\varepsilon}$ by Komlós, Pach, and Woeginger [KPW92].

In geometric applications, this abstract framework is used as follows. The ground set X is typically a set of simple geometric objects (points, lines, hyperplanes), and the ranges in $\mathcal{R}$ are defined in terms of intersection with (or, for point objects, containment in) simply-shaped regions (halfspaces, balls, simplices, etc.), formally assumed to be regions of *constant descriptive complexity*, meaning that they are semi-algebraic sets defined in terms of a constant number of polynomial equations and inequalities of constant maximum degree. It is known that in such cases the resulting range space $(X, \mathcal{R})$ does have finite VC-dimension (see, e.g., [SA95]).

For example, the main result of our paper concerns the range space in which the objects are points in the plane and the ranges are axis-parallel rectangles; more precisely, each range is the intersection of the ground set with such a rectangle. The *dual* range space in this case is one in which the objects are rectangles and each point p in the plane defines a range which is the subset of the given rectangles that contain p. An ε-net in this case is a subset of the rectangles that covers all the "deep" points.

One of the major questions in the theory of ε-nets, open since their introduction more than 20 years ago, is whether the factor $\log \frac{1}{\varepsilon}$ in the upper bound on their size is really necessary, especially in typical low-dimensional geometric situations. To be precise, in the general abstract context the answer is "yes", as shown by Komlós, Pach, and Woeginger [KPW92], using a randomized construction on abstract hypergraphs (see also [PA95]). However, there is no known lower bound, better than the trivial $\Omega(1/\varepsilon)$, in any "concrete" case, certainly in any geometric situation of the kind mentioned above. The prevailing conjecture is that, at least in these geometric scenarios, there always exists an ε-net of size $O(1/\varepsilon)$ [MSW90].

This "linear" upper bound has indeed been established for a few special cases, such as point objects and halfspace ranges in two and three dimensions, and point objects and disk or pseudo-disk ranges in the plane; see [MSW90, Mat92b, CV07, HKSS08, PR08]. Additional progress was made recently. Clarkson and Varadarajan [CV07], essentially adapting Matoušek's technique [Mat92b] to their more general setting, have introduced a method for constructing small-size ε-nets in dual range spaces arising in geometric situations where, as above, the ground set is a collection of regions, and each point p determines a range equal to the set of those regions which contain p, and where the combinatorial complexity of the *union* of any finite number r of the regions in the ground set is small, specifically $o(r \log r)$. (The exact condition is slightly more involved—see below.) As a matter of fact, albeit not explicitly presented in this manner, the technique of [CV07] is more general and can also be applied to the primal version of the problem, provided that it satisfies a condition analogous to the one on small union complexity; see below for more details. More recently, Pyrga and Ray [PR08] have proposed a general abstract scheme for constructing small-size ε-nets in hypergraphs (i.e., range spaces) which satisfy certain properties, and have applied it to the special cases of halfspaces in two and three dimensions, and to several other related scenarios.

The set cover and hitting set problems. Given a range space $(P, \mathcal{R})$, with P and $\mathcal{R}$ finite, the SET COVER problem is to find a minimum-size subcollection $S \subseteq \mathcal{R}$, whose union covers P. A related (dual) problem is the HITTING SET problem, where we want to find a smallest-cardinality subset $H \subseteq P$, with the property that each range $r \in \mathcal{R}$ intersects H. Equivalently, a set cover for $(P, \mathcal{R})$ is a hitting set for the dual range space. The general (primal and dual) problems are NP-hard to solve (even approximately) [GJ79, Kar72], and the simple greedy algorithm yields the (asymptotically) best known approximation factor of $O(1 + \log |P|)$ computable by a polynomial-time algorithm [BGLR93, Fei98]. Most of these problems remain NP-hard even in geometric settings [FG88, FPT81]. However one can attain an improved approximation factor of $O(\log \text{OPT})$ in polynomial time for many of these scenarios, where OPT is the size of the optimal solution. This improvement is based on the technique of Brönnimann and Goodrich [BG95] (see also Clarkson [Cla93]), where the key observation is the relation to ε-nets: The existence of an ε-net of size $O\left(\frac{1}{\varepsilon}\varphi\left(\frac{1}{\varepsilon}\right)\right)$, for any $\varepsilon > 0$, implies that the Brönnimann–Goodrich technique generates, in expected polynomial time, a hitting set (or a set cover) whose size is $O(\text{OPT} \cdot \varphi(\text{OPT}))$.

Hence, for range spaces of finite VC-dimension, the Haussler–Welzl theorem leads to an approximation factor $O(\log \text{OPT})$. Consequently, improved bounds for the size of ε-nets, in the primal or the dual setting, imply improved approximation factors for the corresponding HITTING SET or SET COVER problems, at least in the context of randomized polynomial-time construction (which is what the Brönnimann–Goodrich procedure provides).

Our results. In this paper we first consider the cases of point objects and axis-parallel rectangular ranges in the plane, and of point objects and axis-parallel box ranges in three dimensions, and show that both range spaces admit ε-nets of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, thus significantly improving the standard bound $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. Our technique is similar in spirit to those of Chazelle and Friedman [CF90] and of Clarkson and Varadarajan [CV07], but it differs from them in one key (and fairly simple) idea, which, incidentally, can also be used in the more general context of [CV07] to improve the bounds that are obtained there for the size of the respective ε-nets—see below. An interesting feature of our technique is that it can be extended to points and axis-parallel boxes in *any* dimension, provided that the input points are randomly and uniformly distributed in the unit cube.

We also describe how to construct these ε-nets in randomized expected nearly-linear time. Our results then lead to randomized polynomial-time approximation algorithms for the HITTING SET problem in these two range spaces, involving axis-parallel rectangles and boxes, respectively, which guarantee an approximation factor of $O(\log\log \text{OPT})$.

We then extend our technique to the case of planar point sets and α-*fat* triangles, that is, triangles, each of whose angles is at least α, for some constant $\alpha > 0$ (see [MPSSW94]). In this case too we show the existence of ε-nets of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, leading to an approximation factor of $O(\log\log \textsc{Opt})$ for the corresponding HITTING SET problem.

Similarly, we obtain improved bounds for the size of ε-nets in the dual range space, and, consequently, for approximation factors for the corresponding SET COVER problem, in the following cases, all involving points and regions in the plane (refer to Figure 10):

- *α-fat triangles.* In this case the size of the corresponding ε-net is $O\left(\frac{1}{\varepsilon}\log\log\log\frac{1}{\varepsilon}\right)$, and, as a consequence, the approximation factor for the SET COVER problem becomes $O(\log\log\log \textsc{Opt})$.
- *Locally γ-fat objects*, that is, objects o satisfying the property that, for any disk D whose center lies in o, such that D does not fully contain o in its interior, we have $\text{area}(D \sqcap o) \geq \gamma \cdot \text{area}(D)$, where $D \sqcap o$ is the connected component of $D \cap o$ that contains the center of D (see [dB08]). If we also assume that the boundary of each object has only $O(1)$ locally x-extreme points, and the boundaries of any pair of input objects intersect in at most s points, for some constant s, then the size of the ε-net is $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, and the approximation factor for the SET COVER problem is $O(\log\log \textsc{Opt})$.
- *Locally γ-fat objects of (roughly) equal sizes.* Assuming that the objects satisfy the conditions in the previous case, and that the diameters of any pair of objects differ by at most some constant ratio, the bound on the size of the ε-net improves to $O\left(\frac{1}{\varepsilon}\log\beta_{s+2}(\frac{1}{\varepsilon})\right)$, where $\beta_t(q) := \lambda_t(q)/q$, and $\lambda_t(q)$ is the (nearly linear) maximum length of Davenport-Schinzel sequences of order t on q symbols (see [SA95]). The corresponding approximation factor becomes $O(\log\beta_{s+2}(\textsc{Opt}))$ (see Section 5 for a more detailed discussion of these bounds).
- *Semi-unbounded pseudo-trapezoids*, each consisting of all points lying above some x-monotone arc (or all points lying below such an arc), each pair of which meet at most s times, for s a constant; see Section 5 for a precise definition. In this case the size of the ε-net is, as in the preceding case, $O\left(\frac{1}{\varepsilon}\log\beta_{s+2}(\frac{1}{\varepsilon})\right)$ and the approximation factor is $O(\log\beta_{s+2}(\textsc{Opt}))$. If the pseudo-trapezoids are also unbounded in the x-direction (so they become "pseudo-halfplanes") these bounds slightly improve to $O\left(\frac{1}{\varepsilon}\log\beta_s(\frac{1}{\varepsilon})\right)$ and $O(\log\beta_s(\textsc{Opt}))$, respectively.
- *Jordan arcs with three intersections per pair*, where each of the actual objects is the region bounded by some Jordan arc which starts and ends on the x-axis (and otherwise lies above it) and by the portion of the x-axis between these endpoints, and each pair of the bounding Jordan arcs intersect at most three times. In this case, assuming that none of the given objects "wiggles" too much (as in the case of locally γ-fat objects), the size of the ε-net is $O\left(\frac{1}{\varepsilon}\log\alpha(\frac{1}{\varepsilon})\right)$, and the approximation factor is $O(\log\alpha(\textsc{Opt}))$, where $\alpha(\cdot)$ is the (extremely slowly growing) inverse Ackermann function.

Our technique for rectangles—a brief overview. We start with a brief overview of our analysis, in which we assume some familiarity with the earlier papers [CF90, CV07] cited above. Let P be a given set of n points in the plane. We first sketch a somewhat simpler approach that *almost* works—it does not properly address a certain critical technical issue, but captures the essence of our method. We then briefly describe how to modify it so that it does produce ε-nets of the desired size.

Put $r = 2/\varepsilon$. We draw a random sample R of $s \gg r$ points of P (the specific choice of s, made below, is crucial), and make R part of the ε-net to be constructed, so we only need to handle axis-parallel rectangles which contain at least n/r points, but are *R-empty*, i.e., (axis-parallel) rectangles

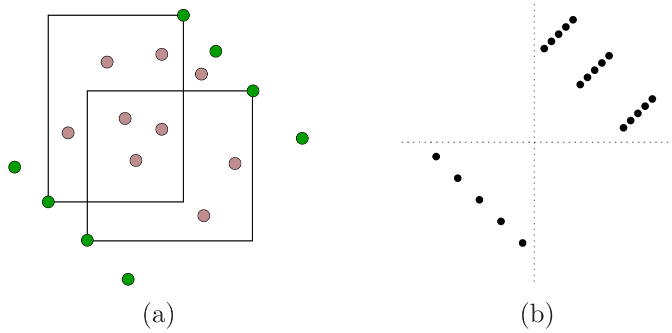


Figure 1: (a) A configuration with quadratically many maximal R-empty rectangles (the points of R are shaded darker and lie on the two extreme staircases). (b) A configuration with an expected quadratic number of maximal R-empty rectangles, each containing $\Omega(n/s)$ points. The lower staircase contains $n/2$ points, and each of the s upper "diagonals" contains $\frac{n}{2s}$ points.

which do not contain any point of R. To "pierce" every such rectangle, we form the subset $\mathcal{M}$ of maximal R-empty rectangles, so that any other R-empty rectangle is contained in one of them. By the standard ε-net theory of [HW87], with high probability each rectangle of $\mathcal{M}$ contains at most $O\left(\frac{n}{s}\log s\right)$ points of P. Moreover, in a sense that we do not make very precise here, the expected number of points of P in such a rectangle is $O(n/s)$. Since $s \gg r$, most rectangles of $\mathcal{M}$ contain fewer than $\varepsilon n = n/r$ points of P, so an R-empty rectangle Q with at least n/r points will not fit into any of them, and we can simply ignore them. For each of the relatively few "heavy" rectangles M of $\mathcal{M}$, we apply the resampling technique of [CF90, CV07], and sample a small subset of $O(t \log t)$ points of $M \cap P$, where $t = s|M \cap P|/n$, to serve as a $(1/t)$-net for $M \cap P$. The union of R and all these samples constitutes the desired ε-net; it is fairly easy to show that this is indeed an ε-net.

This approach does not quite work, because, for a bad choice of R, the number of maximal R-empty rectangles can be $\Theta(s^2)$ in the worst case (see, e.g., [NLH84] and Figure 1(a)). Moreover, even if we only consider random subsets R, there are point sets where the *expected* number of maximal R-empty rectangles which contain $\Omega(n/s)$ points of P is still $\Theta(s^2)$; see Figure 1(b). Using the technique outlined above literally, turns out to yield a bound of $\Theta\left(\frac{1}{\varepsilon^2}\right)$ on the expected size of the ε-net in the worst case, which is of course much too large.

We overcome this issue by modifying the scheme, so that it produces fewer maximal empty rectangles. To do so, we decompose the plane into a binary-tree-like hierarchy of vertical strips. For any rectangle $\tilde{Q}$ which contains at least εn points of P, we find the first (highest in the hierarchy) strip-bounding line which crosses $\tilde{Q}$, take one of its halves, Q, which contains at least $\varepsilon n/2 = n/r$ points, and consider only such rectangles in the construction of our net. We thus face subproblems, each involving a vertical strip σ and the corresponding subset $P \cap \sigma$ of P, and ranges which are rectangles that are "anchored" at a specific side of σ (so that they effectively behave like 3-sided unbounded rectangles for $P \cap \sigma$; refer to Figure 2). The number of maximal R-empty rectangles of this type, within σ, is only *linear* in $|R \cap \sigma|$, leading to an overall collection $\mathcal{M}$ of maximal R-empty rectangles of the new kind, whose size is only $O(s \log r)$.

We now choose $s := cr \log\log r$. Using the so-called Exponential Decay Lemma of [AMS98, CF90], one can show that the expected number of maximal heavy empty rectangles that can contain rectangles Q of the above kind is only *sublinear* in r, which in turn implies that the expected size of the ε-net is dominated by the expected size of R, namely, $O(r \log\log r) = O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$.

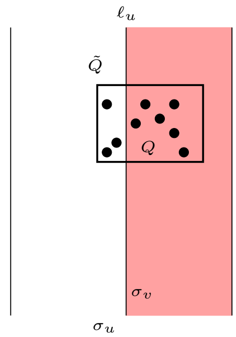


Figure 2: The half-rectangle Q is anchored at the left entry side ℓ_u of the strip σ_v.

Improving the general bounds in [CV07]. Readers familiar with the technique of Clarkson and Varadarajan [CV07] will notice the similarity of our approach to theirs. The key new ingredient is that we use a larger initial sample R, of expected size $\Theta(r \log\log r)$ rather than $O(r)$. The same idea can be applied in the more general context of [CV07], and leads to an improvement of each of their bounds that are super-linear in r. Specifically, Clarkson and Varadarajan consider dual range spaces, and show that if the union complexity of any m of the ranges (i.e., objects in the dual ground set) is $O(m\varphi(m))$, for an appropriate slowly increasing function φ, then there exist ε-nets in such a dual range space of size $O(\frac{1}{\varepsilon}\varphi(\frac{1}{\varepsilon}))$. Using our approach, we obtain ε-nets of size $O(\frac{1}{\varepsilon}\log\varphi(1/\varepsilon))$. Moreover, their method yields improved bounds for ε-nets only when $\varphi(m) = o(\log m)$, whereas our method yields improved bounds as long as $\varphi(m) = 2^{o(\log m)}$. The case of rectangles is interesting in this aspect, because, with the addition of the divide-and-conquer decomposition scheme mentioned above, the complexity of the appropriate analog of the union of m dual ranges (which is the number of maximal empty rectangles) is $O(m \log m)$, which is the threshold bound at which the more "naive" approach of [CV07] fails.[1]

We have just learned that, very recently, Varadarajan [Var08] has independently obtained a similar improvement on the bound of [CV07] for the size of an ε-net in the dual range space of α-fat triangles and planar point sets, using very different methods.

2 Small-size ε-nets for axis-parallel rectangles

Let P be a set of n points in the plane. Put $r := 2/\varepsilon$ and $s := cr\log\log r$, where $c > 1$ is an arbitrary constant. Construct a balanced binary tree $\mathcal{T}$ over the points of P in their x-order, and terminate the tree at the level where the size of each leaf node is between n/r and $n/(2r)$. By construction, $\mathcal{T}$ has at most $1 + \log r$ levels.

Fix a random sample $R \subseteq P$, so that each point $p \in P$ is chosen independently to be included in R with probability $\pi := s/n$; thus the expected size of R is s. The sample R is part of the ε-net N that we are about to construct.

Each node v of $\mathcal{T}$ is associated with a subset P_v of P (resp., R_v of R), consisting of those points of P (resp., of R) stored at the subtree rooted at v. We also associate with v a vertical line ℓ_v which splits P_v into the two subsets P_{v_1}, P_{v_2} associated with the children v_1, v_2 of v. Using the lines ℓ_u, we associate with each node v a strip σ_v, which contains P_v (and R_v), where σ_{root} is the entire plane, and, for a left (resp., right) child node $v \neq$ root of its parent u, σ_v is the left (resp., right) portion of σ_u delimited by ℓ_u. We call ℓ_u the *entry side* of σ_v.

[1]As already noted above, the $\log m$ factor comes from the binary-tree hierarchy—see below for details.

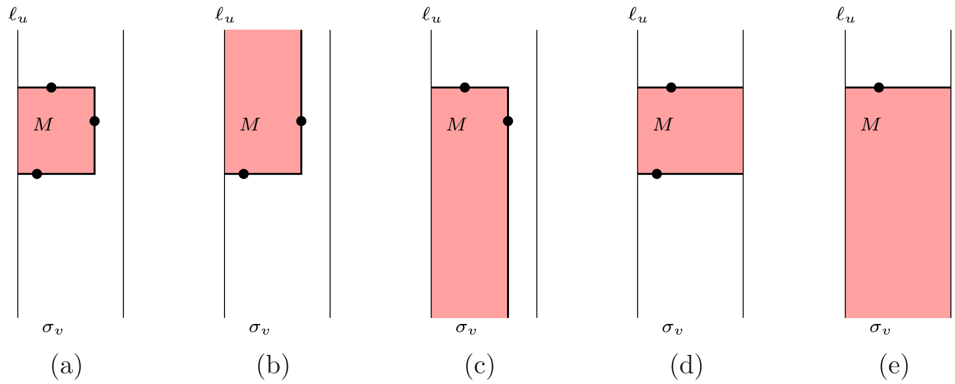


Figure 3: An anchored maximal R-empty rectangle that is determined by three points (a), by a pair of points (b)–(d), or by a single point (e).

Note that, since the sets P_v are defined ahead of the draw of R, our sampling model guarantees that, for each node v, R_v is an unbiased sample of P_v, drawn from P_v by exactly the same rule, namely, by choosing each point independently with probability π.

Let $\tilde{Q}$ be an axis-parallel rectangle containing at least εn points of P, and let u be the highest node of $\mathcal{T}$ such that ℓ_u crosses $\tilde{Q}$, partitioning it into two parts, one of which necessarily contains at least $\varepsilon n/2 = n/r$ points of P. Denote that portion of $\tilde{Q}$ by Q, and let v be the child of u such that $Q \subseteq \sigma_v$. We say that Q is *anchored* at the entry side ℓ_u of σ_v; see Figure 2.

If Q contains a point of R, we are done, as $Q \subset \tilde{Q}$ and the goal was to construct a subset of P that meets every rectangle $\tilde{Q}$ containing at least εn points of P. So we may assume that Q does not contain such a point; we then say that Q is *R-empty*; equivalently, Q is *R_v-empty*.

We define, for each node v of $\mathcal{T}$, a set $\mathcal{M}_v$ consisting of all the maximal (open) anchored R_v-empty axis-parallel rectangles contained in σ_v. Without loss of generality, assume that the entry side ℓ_u of σ_v is its left side. In general, a rectangle M in $\mathcal{M}_v$ is determined by three points of R_v, one point lying on each of the three unanchored sides of M (see Figure 3(a)), but $\mathcal{M}_v$ may also contain degenerate rectangles M where some (or all) of these points are missing, in which case M extends as much as possible, within σ_v, in the appropriate direction (upwards, downwards, or to the right). In particular, when $R_v = \emptyset$, there is precisely one maximal R_v-empty rectangle, namely the whole strip; see Figure 3(b)–(e), illustrating some of these cases.

It is easy to show that $|\mathcal{M}_v| = 2r_v + 1$, where $r_v := |R_v|$. Indeed, if a rectangle M has a point $q \in R_v$ on its right side, then q cannot lie on the right side of any other rectangle in $\mathcal{M}_v$, so the number of such rectangles is r_v (equality is also easy to verify). Otherwise, the points of R_v on the top and bottom sides of M must be consecutive in R_v in the y-order, and there are $r_v - 1$ such pairs. Finally, there are two semi-unbounded rectangles, one delimited from below by the highest point of R_v, and the other delimited from above by the lowest point (as in Figure 3(e)). It is easily checked that the bound $2r_v + 1$ also applies when $r_v = 0, 1$. It thus follows that the overall number of such maximal empty rectangles $M \in \mathcal{M}_v$, over all nodes v of $\mathcal{T}$ at any fixed level, is $O(|R| + r')$, where r' is the number of nodes at the level, and the total over all levels of $\mathcal{T}$ is $O(r + |R| \log r)$.

Returning now to the anchored rectangle Q and the corresponding node v, we note that Q is contained in at least one rectangle in $\mathcal{M}_v$. Indeed, assuming, as above, that the entry side of σ_v is its left side, expand Q by pushing its right side to the right until it touches a point of R_v or reaches the right side of σ_v, and then push the top and bottom sides until each of them meets a point of R_v or extends to $\pm\infty$. The resulting rectangle belongs to $\mathcal{M}_v$ and encloses Q.

For each node v of $\mathcal{T}$, and each member $M \in \mathcal{M}_v$, define the *weight factor* t_M of M to be $s|M \cap P|/n$. Rectangles M with $t_M < s/r = c \log\log r$ can be ignored, because they contain fewer than n/r points of P, so no anchored rectangle Q, as above, can be completely contained in one of them. By the standard ε-net theory [HW87], for each $M \in \mathcal{M}_v$ with $t_M \geq c \log\log r$, there exists a subset $N_M \subseteq M \cap P_v$ of size $c' t_M \log t_M$ that forms a $(1/t_M)$-net for $M \cap P_v$, where c' is another absolute constant.

The final ε-net N is the union of R with the sets N_M, over all the heavy rectangles M (i.e., rectangles with $t_M \geq c \log\log r$) in the respective sets $\mathcal{M}_v$, over all nodes v of $\mathcal{T}$.

N **is an ε-net.** Since $R \subseteq N$, it suffices to show that for any R-empty rectangle Q, contained in a strip σ_v, anchored at the entry side of σ_v, and containing at least $\varepsilon n/2 = n/r$ points of P (i.e., of P_v), and for any $M \in \mathcal{M}_v$ which contains Q, we have $Q \cap N_M \neq \emptyset$. We have

$$\frac{|Q \cap P|}{|M \cap P|} \geq \frac{n/r}{n t_M / s} = \frac{c \log\log r}{t_M} \geq \frac{1}{t_M}.$$

Since N_M is a $(1/t_M)$-net for $M \cap P$, it follows that $Q \cap N_M \neq \emptyset$, as asserted. Note that the above inequality implies that we need not sample that many points in N_M, and can make do with $c' t_M^* \log t_M^*$ points, where $t_M^* := t_M/(c \log\log r)$. However, this slight improvement does not asymptotically affect the bound that we are about to derive.

Estimating the size of N. The expected size of N is equal to

$$\mathbf{Exp}\left\{|R| + c' \sum_v \sum_{\substack{M \in \mathcal{M}_v \\ t_M \geq c \log\log r}} t_M \log t_M \right\} = cr \log\log r + c' \cdot \mathbf{Exp}\left\{\sum_v \sum_{\substack{M \in \mathcal{M}_v \\ t_M \geq c \log\log r}} t_M \log t_M \right\}.$$

We continue the analysis using the notation of [AMS98]. Fix a level i; each node v at this level satisfies $|P_v| = n/2^i$. Let $\mathrm{CT}(R)$ denote the union of the collections $\mathcal{M}_v$, over all nodes v at level i. For a positive parameter t, let $\mathrm{CT}_t(R)$ denote the subset of $\mathrm{CT}(R)$ consisting of those rectangles M with $t_M \geq t$. Let R' denote another random sample of P, where each point $p \in P$ is now chosen, independently, to belong to R' with probability $\pi' := \pi/t$.

Let $\mathcal{C}$ denote the set of all rectangles M, such that M is anchored at the entry side of σ_v, for some node v at level i, and has one point of P on each of its three other sides (the cases of degenerate rectangles, determined by fewer than three points, are treated in a fully analogous manner). For a rectangle $M \in \mathcal{C}$, its *defining set* $D(M)$ is the set of these three points, and its *killing set* $K(M)$ is the set of points of P in the interior of M. (Recall that throughout this discussion we have fixed the level i.)

Agarwal *et al.* [AMS98] impose two axioms on the sets $\mathrm{CT}(R)$. These axioms are too intricate for what we need here, while they are necessary to handle the more involved scenario considered in [AMS98]. For our purpose, we can replace them by the single "axiom," asserting that a rectangle $M \in \mathcal{C}$ belongs to $\mathrm{CT}(R)$ if and only if $D(M) \subseteq R$ and $K(M) \cap R = \emptyset$, which holds by construction in our setting. (We also caution the reader that our sampling model is different from that of [AMS98]—they sample a random subset of a fixed given size uniformly from all such subsets, whereas we independently choose each point of P to belong to the sample. Nevertheless, the lemma, given below, also holds in our model; if at all, the analysis is simpler. For the sake of completeness, we give, in the appendix, a short (but complete) proof of our variant of the lemma.)

Lemma 2.1 (Exponential Decay Lemma; Agarwal *et al.* [AMS98])**.**

$$\mathbf{Exp}\left\{|\,\mathrm{CT}_t(R)|\right\} = O\left(2^{-t}\, \mathbf{Exp}\left\{|\,\mathrm{CT}(R')|\right\}\right).$$

We apply the lemma with $t = c \log\log r$, so $\pi' = \pi/t = r/n$. Recall that $\mathrm{CT}(R')$ is the set of all maximal R'-empty rectangles, anchored at the entry sides of their respective strips σ_v at the fixed level i. Their number is $|\mathrm{CT}(R')| = \sum_v (2r'_v + 1)$, where $R'_v := R' \cap \sigma_v$, and $r'_v := |R'_v|$. Since the sets R'_v at level i are disjoint, $\sum_v r'_v = |R'|$. Hence, since there are at most $2r$ nodes at a fixed level of the tree, we have $|\mathrm{CT}(R')| \le 2|R'| + 2r$. Hence $\mathbf{Exp}\left\{|\mathrm{CT}(R')|\right\} = O(r)$. We thus have

$$\mathbf{Exp}\left\{|\mathrm{CT}_t(R)|\right\} = O\left(2^{-t}\,\mathbf{Exp}\left\{|\mathrm{CT}(R')|\right\}\right) = O\left(r2^{-c\log\log r}\right) = O\left(r/\log^c r\right).$$

More generally, for any $j \ge t$, we have $\mathbf{Exp}\left\{|\mathrm{CT}_j(R)|\right\} = O(r/2^j)$, as is easily checked.

Getting back to the contribution of the fixed level i to the expected size of N, we have (where $t = c\log\log r$)

$$\begin{aligned}
\mathbf{Exp}\Bigg\{\sum_{v \text{ at level } i}\ \sum_{\substack{M\in\mathcal{M}_v\\ t_M\ge t}} t_M \log t_M\Bigg\} &= \mathbf{Exp}\Bigg\{\sum_{j\ge t}\ \sum_{\substack{M\in\mathrm{CT}(R)\\ t_M=j}} j\log j\Bigg\} \qquad (*)\\
&= \mathbf{Exp}\Bigg\{\sum_{j\ge t} j\log j\cdot\left(|\mathrm{CT}_j(R)| - |\mathrm{CT}_{j+1}(R)|\right)\Bigg\}\\
&= \mathbf{Exp}\Bigg\{t\log t\cdot|\mathrm{CT}_t(R)|\\
&\qquad + \sum_{j>t}\big(j\log j - (j-1)\log(j-1)\big)|\mathrm{CT}_j(R)|\Bigg\}\\
&= O\left(\frac{r}{\log^c r}(t\log t) + \sum_{j>t}\frac{r}{2^j}\log j\right)\\
&= O\left(\frac{rt\log t}{\log^c r}\right) = O\left(\frac{r\log\log r\log\log\log r}{\log^c r}\right).
\end{aligned}$$

Recall again that the analysis so far has been confined to a single level i. Repeating it for each of the $1 + \log r$ levels, we obtain, recalling that $c > 1$,

$$\mathbf{Exp}\left\{|N|\right\} = O\left(r\log\log r + \frac{r\log\log r\log\log\log r}{\log^{c-1} r}\right) = O(r\log\log r).$$

We have thus shown

Theorem 2.2. *For any set P of n points in the plane and a parameter $\varepsilon > 0$, there exists an ε-net of P, of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, for axis-parallel rectangles.*

Remark: A key ingredient of the analysis is that we have managed to reduce the expected number of R-empty rectangles from $\Theta(s^2)$ to $O(s\log r)$, using a decomposition of the point set into canonical subsets, so that (i) any rectangle $\tilde{Q}$ with at least εn points of P interacts with just *two* subsets (any constant number would do just as well), and (ii) for each canonical subset, the number of maximal R-empty rectangles (now anchored at the entry side of the respective strip and fully contained in that strip) is only linear in the number of sample points in that strip.

Constructing the ε-net. We next present a randomized algorithm for constructing an ε-net of the above size.

We construct the balanced binary tree $\mathcal{T}$ over the points of P in $O(n \log r)$ time (stopping at nodes v for which $n/(2r) < |P_v| \leq n/r$), and generate the random sample R, using the drawing model assumed above; the expected size of R is s.

Following the above notation, we associate with each node $v \neq$ root of $\mathcal{T}$ a strip σ_v, the subsets P_v, R_v, and an entry side ℓ_u of σ_v (where u is the parent of v). We next construct, for each such node v, the set $\mathcal{M}_v$ of all maximal anchored R_v-empty axis-parallel rectangles contained in σ_v. This is easy to do in time $O(r_v \log r_v)$, where $r_v := |R_v|$, as follows. Assume that ℓ_u is the left side of σ_v. Sort the points of R_v by their y-coordinates, and find, for each point q, the lowest point q' which lies above q and to its left. This can be done in linear time, by scanning the points of R_v in decreasing y-order, and by dynamically maintaining the sorted sequence of xy-minima [CLRS01]. Symmetrically, we find, for each point q, the highest point q'' which lies below q and to its left. The resulting triples (q, q', q'') (including degenerate ones) determine r_v of the maximal empty rectangles in $\mathcal{M}_v$. Each of the other $r_v + 1$ rectangles straddles σ_v from left to right and either is delimited by a pair of points of R_v, consecutive in the y-order, which lie on its top and bottom sides, or is an unbounded half-strip, bounded by a single point.

It thus follows that the overall expected running time for constructing the sets $\mathcal{M}_v$, over all nodes v at a fixed level i, is $O(s \log r)$, for a total of $O(s \log^2 r)$ time, over all levels i.

We next count, for each resulting rectangle M, the number of points in $M \cap P$, using a standard 2-dimensional range-tree data structure. This yields the respective weight factors t_M, as defined above; we keep only those rectangles with $t_M \geq c \log\log r$. For each of these surviving rectangles M, we *report* the set $P \cap M$, and construct a $(1/t_M)$-net for $P \cap M$, using, e.g., the deterministic algorithm of Matoušek [Mat95] (or a straightforward random sampling mechanism [HW87]). We output the union of R with all the resulting nets. Using the Exponential Decay Lemma and similar considerations as in the proof of Theorem 2.2, it can be shown that the overall expected number of reported points in the sets $P \cap M$, over all heavy rectangles M and nodes v, is only linear in n.

As argued above, the output N is guaranteed to be an ε-net for P (if we construct the sub-nets N_M deterministically). The size of N is a random variable whose expectation is $O(r \log\log r)$. We can ensure this size with high probability, by discarding outputs that are too large and by repeating the sampling.

The entire algorithm takes $O(n \log n)$ randomized expected time, as is easily seen.

Remark. The running time of the algorithm can be slightly improved to $O(n \log r)$. First, we apply the sampling with $s = cr \log\log r$, where c is substantially larger than 1, say $c = 2$. As already observed, constructing the truncated tree $\mathcal{T}$, with the stopping condition that we use, can be done in $O(n \log r)$ time, and this also holds for the construction of the set $\mathcal{M}$ of all maximal anchored R-empty rectangles. Regarding the range-searching procedure, it is sufficient to *approximate* the count $|M \cap P|$ up to an additive term of $O(n/r)$, with a sufficiently small constant of proportionality. To do so, we construct the range-tree, but stop at primary nodes u whose associated subsets P_u satisfy $n/(2c_1 r) < |P_u| \leq n/(c_1 r)$, for some constant $c_1 > 1$. Similarly, we stop the construction of each secondary tree $\mathcal{T}_u$ at nodes v for which $n/(2c_1 r \log r) < |P_{u,v}| \leq n/(c_1 r \log r)$. It is easily checked that each range-counting query can miss at most $4n/(c_1 r)$ points. We thus find all rectangles whose approximated count is at least $(1 - 4/c_1)n/r$, and report the points in each of them. As above, the overall expected number of reported points is only $O(n)$. We then construct a $(1/t_M)$-net for each of the heavy rectangles M with $t_M \geq c \log\log r$, and continue as in the original algorithm. By choosing c_1 sufficiently large, we can ensure that (a) all canonical empty rectangles that can contain heavy rectangles are reported, and (b) the overall expected size of N remains $O\left(\frac{1}{\varepsilon} \log\log \frac{1}{\varepsilon}\right)$. The overall running time is easily seen to be only $O(n \log r)$.

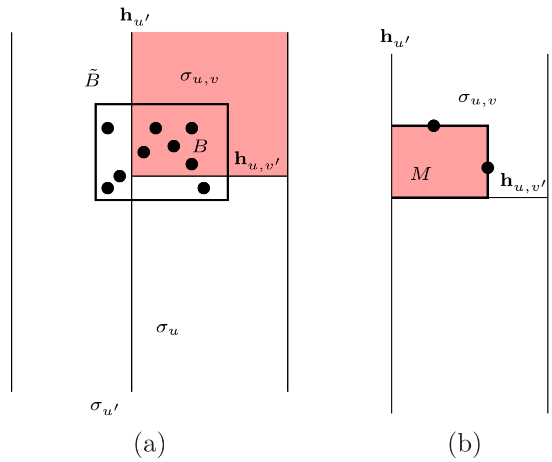


Figure 4: A two-dimensional illustration: (a) The box B is anchored at the (apex of the) quadrant $\sigma_{u,v}$ (octant in 3-space). (b) An anchored box that is determined by a pair of points (a triple in 3-space).

3 Small-size ε-nets for axis-parallel boxes in three dimensions

We next extend our construction to the three-dimensional case. We now let P be a set of n points in $\mathbb{R}^3$, and put $r := 8/\varepsilon$ and $s := cr \log\log r$, for some fixed constant $c > 3$. We use a similar sampling model as in the two-dimensional problem, in order to generate a random subset $R \subseteq P$ of expected size s.

We next construct a three-level range-tree $\mathcal{T}$, over the points of P (see, e.g., [dBCKO08]), where the points are sorted by their x-coordinates in the primary tree, by their y-coordinates in each secondary tree, and by their z-coordinates in each tertiary tree. We associate with each node u of the primary tree the subset P_u of points that it represents, and a secondary (y-sorted) tree $\mathcal{T}_u$ on P_u. Similarly, with each node v of a secondary tree $\mathcal{T}_u$ we associate the corresponding subset $P_{u,v}$ of P_u and a tertiary (z-sorted) tree $\mathcal{T}_{u,v}$. Finally, each node w of a tertiary tree $\mathcal{T}_{u,v}$ is associated with the corresponding subset $P_{u,v,w}$ of $P_{u,v}$. We construct each of the three levels of $\mathcal{T}$ down to nodes for which the size of their associated subset is between n/r and $n/(8r)$. Clearly, each of the primary, secondary, and tertiary trees has at most $3 + \log r$ levels, and the total number of nodes in the range-tree $\mathcal{T}$ is $O(r \log^2 r)$. Moreover, the sum of the sizes of all the subsets stored at the various nodes is $O(n \log^3 r)$; see, e.g., [dBCKO08] for further details.

Following the notation of Section 2, we associate with each non-leaf node of any subtree an axis-parallel plane which evenly splits the subset stored at the node into the two subsets stored at its children. More specifically, each non-leaf node u of the primary tree stores a plane $\mathbf{h}_u$ orthogonal to the x-axis, each non-leaf node v of a secondary tree $\mathcal{T}_u$ stores a plane $\mathbf{h}_{u,v}$ orthogonal to the y-axis, and each non-leaf node w of a tertiary tree $\mathcal{T}_{u,v}$ stores a plane $\mathbf{h}_{u,v,w}$ orthogonal to the z-axis.

These planes define, for each node w of a tertiary tree $\mathcal{T}_{u,v}$, an octant $\sigma_{u,v,w}$ which is the intersection of three halfspaces $H_u \cap H_{u,v} \cap H_{u,v,w}$, where (i) H_u is the halfspace bounded by $\mathbf{h}_{u'}$ and containing P_u, where u' is the parent of u; (ii) $H_{u,v}$ is the halfspace bounded by $\mathbf{h}_{u,v'}$ and containing $P_{u,v}$, where v' is the parent of v in $\mathcal{T}_u$; and (iii) $H_{u,v,w}$ is the halfspace bounded by $\mathbf{h}_{u,v,w'}$ and containing $P_{u,v,w}$, where w' is the parent of w in $\mathcal{T}_{u,v}$. In what follows we only consider triples (u, v, w) of vertices, each of which has a parent in its respective tree. Thus all three halfspaces are proper, and $\sigma_{u,v,w}$ is a non-degenerate octant. (Note, though, that, in general, it is more accurate to regard $\sigma_{u,v,w}$ as a box, or a clipped octant, bounded on the other side also by planes associated with ancestors of u, v, and w. Nevertheless, in most of the following analysis, it suffices to treat $\sigma_{u,v,w}$ as an octant.)

Let B_0 be an axis-parallel box containing at least εn points of P. Let u' be the highest node in $\mathcal{T}$, so that the plane $\mathbf{h}_{u'}$ meets B_0. This plane partitions B_0 into two portions, one of which, call it B_1, contains at least $\varepsilon n/2$ points of P. Let u be the corresponding child of u' so that H_u contains B_1. Next, let v' be the highest node in $\mathcal{T}_u$, such that $\mathbf{h}_{u,v'}$ meets B_1, partitioning it into two portions, one of which, B_2, contains at least $\varepsilon n/4$ points of P. Let v be the child of v' for which $H_{u,v'}$ contains B_2. Finally, let w' be the highest node in $\mathcal{T}_{u,v}$, such that $\mathbf{h}_{u,v,w'}$ meets B_2, partitioning it into two portions, one of which, B, contains at least $\varepsilon n/8$ points of P. Let w be the child of w' for which $H_{u,v,w'}$ contains B. (Note that u, v, w are well defined, in the sense that each of the sub-boxes is indeed split by a plane associated with a node in the corresponding truncated tree, and does not reach a leaf without being split.)

By construction, B is *anchored* at the resulting octant $\sigma := \sigma_{u,v,w}$, in the sense that the apex o of σ is a vertex of B, and the three facets of B adjacent to o lie on the three respective axis-parallel planar quadrants bounding σ. Moreover, as far as the set $P_{u,v,w}$ is concerned, we can replace B by an octant which is oppositely oriented to σ, and whose apex is the vertex o' of B opposite to o. See Figure 4(a) for an illustration of (the 2-dimensional analog of) this scenario.

For each node w of a tertiary tree $\mathcal{T}_{u,v}$, put $R_{u,v,w} = R \cap \bar{\sigma}_{u,v,w}$, where $\bar{\sigma}_{u,v,w}$ is the actual box that the "octant" $\sigma_{u,v,w}$ represents (see the comment above), and $r_{u,v,w} = |R_{u,v,w}|$. Let $\mathcal{M}_{u,v,w}$ denote the set of all maximal anchored R-empty (i.e., $R_{u,v,w}$-empty) axis-parallel boxes contained in the octant $\sigma_{u,v,w}$. Since each box $M \in \mathcal{M}_{u,v,w}$ behaves as an octant inside $\sigma_{u,v,w}$, it is determined by at most three points of $R_{u,v,w}$, each lying on a distinct facet of M; see Figure 4(b) for a two-dimensional illustration. The number of such empty boxes (or, rather, octants) is only $O(r_{u,v,w}+1)$, as shown[2] in [BSTY98, KRSV07]. It thus follows that the overall size of the sets $\mathcal{M}_{u,v,w}$, over all nodes w of all tertiary trees $\mathcal{T}_{u,v}$, is $O(|R| \log^3 r + r \log^2 r)$.

We proceed as in the planar case. We make R part of the output ε-net, thereby disposing of any box B_0 whose resulting anchored portion B contains a point of R. For any other box B_0, the corresponding portion B is R-empty, and it is then easy to show that B is contained in at least one maximal R-empty box M in the set $\mathcal{M}_{u,v,w}$ of the corresponding octant $\sigma_{u,v,w}$. Moreover, the weight factor t_M of M, defined as in the planar case, must satisfy $t_M \geq c \log \log r$.

Thus, for each such heavy maximal box M, we take a $(1/t_M)$-net N_M, for the set $P \cap M$, of size $O(t_M \log t_M)$, whose existence is guaranteed by [HW87], and output the union N of R with all the resulting nets N_M. Arguing as in the planar case, it is easy to show that N is indeed an ε-net for P.

We bound the expected size of N using similar analysis steps to those in the planar problem. We define $\mathrm{CT}(R)$ to be the union of all the collections $\mathcal{M}_{u,v,w}$, over all nodes w of all tertiary trees $\mathcal{T}_{u,v}$, appearing in a fixed triple of levels i_1 (primary), i_2 (secondary), and i_3 (tertiary). As before, $\mathrm{CT}_t(R)$ is the subset of $\mathrm{CT}(R)$ consisting of those boxes M with $t_M \geq t$, for any parameter t. It is easy to verify that the Exponential Decay Lemma holds in this scenario as well, and thus

$$\mathbf{Exp}\left\{|\mathrm{CT}_t(R)|\right\} = O\left(2^{-t}\, \mathbf{Exp}\left\{|\mathrm{CT}(R')|\right\}\right),$$

where R' is another smaller random sample defined as in Section 2. Next, arguing as in the planar

[2]In fact, the result in [KRSV07] is more general. It asserts that the number of maximal empty orthants for a set of m points in $\mathbb{R}^d$ is $O(m^{\lfloor d/2 \rfloor})$. It is the non-linearity of this bound for $d \geq 4$ which hampers the extension of our technique to higher dimensions.

problem, we obtain that

$$\mathbf{Exp}\left\{ \sum_{v \text{ at levels } i_1,\, i_2,\, i_3} \sum_{\substack{M \in \mathcal{M}_v \\ t_M \geq c \log\log r}} t_M \log t_M \right\} = O\left(\frac{r \log\log r \log\log\log r}{\log^c r} \right).$$

Repeating the analysis for each of the $O\left(\log^3 r\right)$ triples (i_1, i_2, i_3), we obtain that the expectation of the above sum is $o(r)$, provided $c > 3$, as we indeed assume; thus

$$\mathbf{Exp}\left\{|N|\right\} = \mathbf{Exp}\left\{|R|\right\} + o(r) = O(r \log\log r).$$

We have thus shown:

Theorem 3.1. *For any set P of n points in $\mathbb{R}^3$ and a parameter $\varepsilon > 0$, there exists an ε-net of P, for axis-parallel boxes, of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$.*

Constructing the ε-net. We construct an ε-net of this size using an easy extension of the algorithm presented in Section 2. We start by building a 3-level range tree over the points of P, using $O(n \log^2 n)$ time and storage. The enumeration of the maximal anchored $R_{u,v,w}$-empty octants in any canonical octant $\sigma_{u,v,w}$ can be performed in $O(|R_{u,v,w}| \log^2 r)$ time, using the algorithm described in [KRSV07]. Using our range tree, we compute the weight factor t_M of each maximal octant M, collect all the heavy octants M (using counting queries), report the corresponding subsets $P \cap M$, and construct, for each such octant M, a $(1/t_M)$-net of size $O(t_M \log t_M)$, for $P \cap M$, using standard techniques as in the two-dimensional case. Omitting the further easy details, we obtain that the expected running time of the algorithm is $O(n \log^2 n)$, as asserted.

As in the planar case, the algorithm can be slightly improved to $O(n \log^3 r)$, using similar refinements.

Random point sets in any dimension. The technique fails in four and higher dimensions, because the number of maximal empty orthants with respect to a set of m points can be $\Theta\left(m^{\lfloor d/2 \rfloor}\right)$ (see [BSTY98, KRSV07]), which is at least quadratic for $d \geq 4$. It is a challenging open problem to extend our results to four and higher dimensions.

Nevertheless, there is one scenario where the technique works in any dimension, which is the case when the ground set P consists of randomly and uniformly distributed points in $\mathbb{R}^d$. Specifically, we assume that each point of P is chosen independently at random from the uniform distribution in $[0,1]^d$. As shown in [KRSV07], the expected number of maximal empty boxes in this case, for a set of m points, is only $O(m \log^{d-1} m)$ (see also [NLH84] for the planar case). Moreover, our random sampling model (where the random choices are assumed, of course, to be made independently of the random choices made while constructing the input set) ensures that the sample R is also an unbiased set of randomly, independently, and uniformly distributed points, so the expected number of maximal R-empty boxes is $O(s \log^{d-1} s)$ (see Section 5 for a proof of a (more general) bound of this type); the expectation is with respect to both the random drawing of the points of the input set, and our drawing of the sample R.

Since the (expected) number of maximal R-empty boxes is only nearly linear in s, we can carry out the preceding analysis, without having to decompose the input set into canonical strips or orthants, and thus obtain an ε-net of expected size[3] $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$. We have thus shown

[3]By consulting the derivation in (*), it is easily verified that the expected size of the resulting net N is $\mathbf{Exp}\left\{|R|\right\}$ plus a term equal to $\mathbf{Exp}\left\{|CT(R')|\right\}$ divided by a polylogarithmic factor. This implies the bound asserted here.

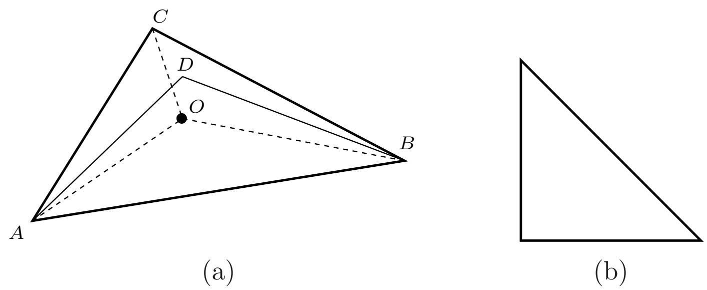


Figure 5: (a) The canonization step. The triangle ABC is covered by three triangles, each of which contains the center O of the inscribed circle of ABC, and has two edge orientations that are taken from a fixed set of directions. Only one of these triangles (ABD) is depicted in the figure. (b) A semi-canonical right triangle, after an appropriate affine transformation.

Theorem 3.2. *For any set P of n points in $\mathbb{R}^d$, each of which is drawn independently from the uniform distribution on $[0,1]^d$, and a parameter $\varepsilon > 0$, there exists an ε-net of P, for axis-parallel boxes, of expected size*[4] $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$.

4 Small-size ε-nets for fat triangles in the plane

In this section we present an extension of our technique to the range space of points in the plane and *α-fat triangles*, for some fixed constant $\alpha > 0$, where a triangle is α-fat if each of its angles is at least α. We thus have a set P of n points in the plane, and a parameter $\varepsilon > 0$, and our goal is to construct a small-size ε-net $N \subseteq P$, so that any α-fat triangle that contains at least εn points of P contains a point of N.

Passing to semi-canonical triangles. Following the analysis of [MPSSW94], we cover each α-fat triangle T by a triple of "semi-canonical" $(\alpha/2)$-fat triangles, each of which has a pair of edges with orientations taken from a fixed finite set $\mathcal{D}$ of $O(1/\alpha)$ directions, and a third edge that bounds T; see [MPSSW94, Lemma 3.2] and Figure 5(a). Clearly, if T contains at least εn points of P then at least one of the three covering triangles contains at least $\varepsilon n/3$ points of P.

This canonization step yields a constant number ($O(1/\alpha^2)$, to be precise) of subfamilies of $(\alpha/2)$-fat triangles, where the triangles in each subfamily have two edges at fixed orientations (in $\mathcal{D}$), and a third edge whose orientation belongs to a sufficiently small range. Our strategy is thus to construct an $(\varepsilon/3)$-net for P and each of these subfamilies, and the union of all these nets will be an ε-net for P and the family of all α-fat triangles.

Thus, in what follows we focus on a fixed semi-canonical family $\mathcal{F}$. As in [MPSSW94], by applying an appropriate affine transformation, we may assume that each triangle $T \in \mathcal{F}$ is a right triangle with one horizontal edge and one vertical edge, which meet at the lower-left vertex of T; see Figure 5(b).

Thus let P and $\mathcal{F}$ be as above, and put $r := 24/\varepsilon$ and $s := cr\log\log r$, for some fixed constant $c > 2$. We use a similar sampling model as in the cases of axis-parallel rectangles and boxes, for drawing a random subset $R \subseteq P$ of expected size s, which becomes part of our ε-net.

We next construct a two-level range-tree $\mathcal{T}$, over the points of P, in an analogous manner to that presented in Section 3. The points are sorted by their x-coordinates in the primary tree, and by their y-coordinates in each secondary tree, and we construct each of the two levels of $\mathcal{T}$ down

[4]The expectation is with respect to the random choice of P.

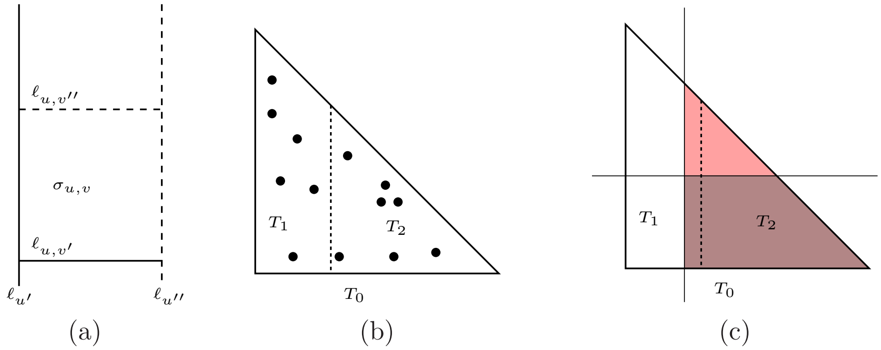


Figure 6: (a) The "quadrant" $\sigma_{u,v}$ is defined by the line splitters $\ell_{u'}$, $\ell_{u,v'}$, but it is also bounded by ancestor splitters $\ell_{u''}$ and $\ell_{u,v''}$. (b) The decomposition of T_0. (c) An anchored triangle T appears as a triangle (lightly shaded) homothetic to T_0 at the upper-right quadrant, or as a right-angle trapezoid (darkly shaded) at the lower-right quadrant.

to nodes for which the size of their associated subset is between n/r and $n/(4r)$. Following the notation of Section 3, each node u of the primary tree is associated with the subset P_u of points that it represents, and a secondary (y-sorted) tree $\mathcal{T}_u$ on P_u, and each node v of any secondary tree $\mathcal{T}_u$ is associated with a corresponding subset $P_{u,v}$ of P_u. Each non-leaf node u of the primary tree stores a vertical line "splitter" ℓ_u, and each non-leaf node v of any secondary tree $\mathcal{T}_u$ stores a horizontal line splitter $\ell_{u,v}$. For each such secondary node v of a tree $\mathcal{T}_u$, the lines $\ell_{u'}$ and $\ell_{u,v'}$, where u' is the parent of u in $\mathcal{T}$ and v' is the parent of v in $\mathcal{T}_u$ (as before, we only handle nodes for which u' and v' exist), define a quadrant $\sigma_{u,v}$, which is the intersection of two halfplanes bounded by $\ell_{u'}$ and $\ell_{u,v'}$ and containing $P_{u,v}$. (Technically, similar to the situation in Section 3, $\sigma_{u,v}$ is a (possibly unbounded) rectangle, where the other vertical and horizontal edges of $\sigma_{u,v}$, if they exist, are portions of respective splitters $\ell_{u''}$, $\ell_{u,v''}$, where u'' is an appropriate ancestor of u' in $\mathcal{T}$ and v'' is an appropriate ancestor of v' in $\mathcal{T}_u$; see Figure 6(a).)

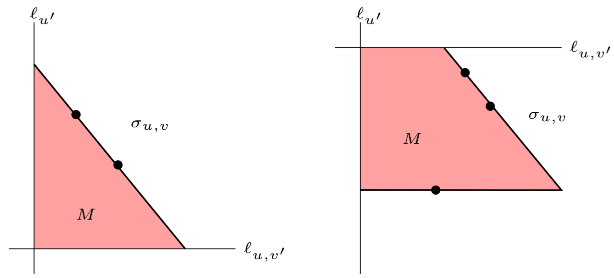
Let T_0 be a right triangle in our semi-canonical family, containing at least $\varepsilon n/3 = 8n/r$ points of P. We first decompose T_0 into two parts, T_1, T_2, by a vertical line, so that T_1 lies to the left of the line and T_2 to its right, and $|T_1 \cap P| \le |T_2 \cap P| \le |T_1 \cap P| + 1$. That is, $|T_1 \cap P| \ge 4n/r - 1$ and $|T_2 \cap P| \ge 4n/r$. See Figure 6(b) for an illustration.

As in the case of axis-parallel boxes, we locate the highest node u' in $\mathcal{T}$, so that the line $\ell_{u'}$ meets T_1, thus splitting T_0 into two parts, where the right part is a triangle T', homothetic to T_0 and *fully containing* T_2. In particular, we have $|T' \cap P| \ge 4n/r$. Let u be the right child of u'. We next locate the highest node v' in $\mathcal{T}_u$, such that $\ell_{u,v'}$ meets T'. We focus on the portion T of T' that contains at least $2n/r$ points, and denote by v the child of v' whose corresponding quadrant $\sigma_{u,v}$ contains T.

Brief discussion. (a) Although it may not appear so at first sight, the analysis just given uses also the fact that $|T_1 \cap P|$ is large, to guarantee the existence of the node u in the primary splitting stage: Since we stop the expansion of the primary tree at nodes containing roughly n/r points each, we need to ensure that T_1 contains sufficiently many points of P, or else it would "fall between the cracks" and not be stabbed by any splitter $\ell_{u'}$. This, however, does not happen due to the way in which T_0 is decomposed,

(b) It is important that T_1 is the portion of T_0 stabbed by $\ell_{u'}$ (and not T_2) because it then ensures that the apex o of $\sigma_{u,v}$ is indeed contained in T_0.

(c) Note that only right children u in the primary tree require the construction of a secondary tree



(a) (b)

Figure 7: A maximal anchored $R_{u,v}$-empty (a) right triangle, and (b) right trapezoid.

$\mathcal{T}_u$.

We now continue with the construction. The clipped region T is either (a) a triangle, homothetic to T_0, whose right-angle vertex is the apex o of $\sigma_{u,v}$, or (b) a right-angle trapezoid, having o as its top-left vertex, so that its bases are horizontal, its left side is vertical, and its right side is a portion of the hypotenuse of T_0; see Figure 6(c). In both cases we refer to T as being *anchored* at o. Note that in case (a) v is a right child of its parent, representing an upper quadrant, and that in case (b) v is a left child. representing a lower quadrant. Also, in both cases the slope of the slanted edge of T is negative, so in case (b) the slanted edge moves "away" from o, making the lower base of T longer than its upper base.

Recall that we have drawn a "global" random sample R of P. For each node v of each secondary tree $\mathcal{T}_u$, we put $R_{u,v} := R \cap \sigma_{u,v}$ and $r_{u,v} = |R_{u,v}|$. We make R part of the output ε-net N, so if T contains a point of R we are done.

To handle the other case, we define a family $\mathcal{M}_{u,v}$ of maximal anchored $R_{u,v}$-empty regions, with the property that each anchored $R_{u,v}$-empty region T (triangle or trapezoid, as above), is covered by at most two regions in $\mathcal{M}_{u,v}$. Each region in $\mathcal{M}_{u,v}$ is either (a) an anchored $R_{u,v}$-empty right triangle whose hypotenuse touches two points of $R_{u,v}$ (that is, it supports an edge of the convex hull of $R_{u,v}$), or (b) an anchored $R_{u,v}$-empty right-angle trapezoid whose slanted side (has negative slope and) touches two points of $R_{u,v}$, and whose unanchored (lower) horizontal base passes through a point of $R_{u,v}$ (which might coincide with one of the two points lying on the slanted edge, i.e., be a vertex of the region), or else lies on the bottom side of the "quadrant" $\sigma_{u,v}$. In each of these cases, the region is clipped within $\sigma_{u,v}$ (when $\sigma_{u,v}$ is defined by three or more splitters). See Figure 7.

In case (a), we also include in $\mathcal{M}_{u,v}$ two axis-parallel rectangles M_1, M_2 anchored at o, so that (i) the right edge of M_1 passes through the leftmost point of $R_{u,v}$ and its top edge lies on the top side of $\sigma_{u,v}$ (if it exists), otherwise M_1 extends to ∞, and (ii) the top edge of M_2 passes through the bottommost point of $R_{u,v}$ and its right edge lies on the right side of $\sigma_{u,v}$ (if it exists), otherwise M_2 extends to ∞. See Figure 8(a). In case (b), we also include in $\mathcal{M}_{u,v}$ axis-parallel rectangles of the following two types: (i) rectangles that are anchored at o, with both right and bottom sides passing through a point of $R_{u,v}$; (ii) rectangles whose left and right sides lie respectively on the left and right sides of $\sigma_{u,v}$ (if the right side exists), and whose top and bottom sides pass through two respective points of $R_{u,v}$, necessarily consecutive in the y-order (including two extreme rectangles, above the highest point and below the lowest point). See Figure 8(b). Finally, if $R_{u,v}$ is empty, $\mathcal{M}_{u,v}$ consists of the single region $\sigma_{u,v}$.

We next claim that $|\mathcal{M}_{u,v}| = O(r_{u,v} + 1)$. This is trivial when $R_{u,v} = \emptyset$, so assume that $R_{u,v}$ is nonempty. The claim is then obvious for regions of type (a), because their number is at most two plus the number of edges of the lower-left convex hull of $R_{u,v}$. To bound the number

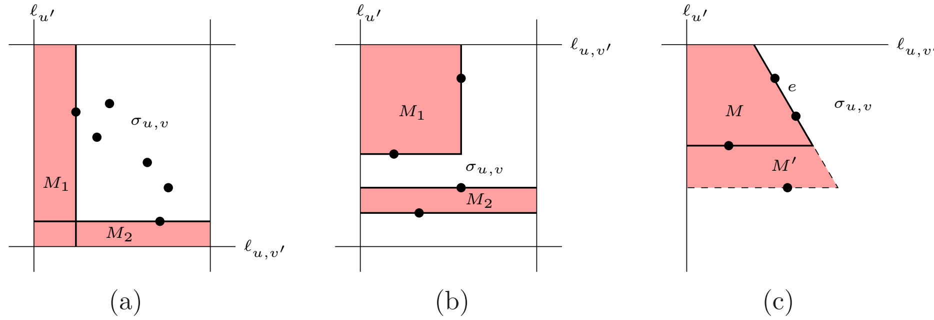


Figure 8: Anchored maximal $R_{u,v}$-empty rectangles (a) for upper quadrants, and (b) for lower quadrants. (c) The larger right trapezoid M' cannot be empty, if it shares its slanted edge e with M.

of regions of type (b), sort the points of $R_{u,v}$ in decreasing y-order, and let the sorted sequence be $(q_1, q_2, \ldots, q_{r_{u,v}})$. Put $R^{(j)} = \{q_1, \ldots, q_{j-1}\}$, for $j = 1, \ldots, r_{u,v}$. Let M be a region of type (b) whose lower horizontal base passes through q_j, so that q_j is not a vertex of M. Then its slanted edge must contain an edge e of the (lower-left) convex hull of $R^{(j)}$. Moreover, if such an M exists then there cannot exist another region M' whose slanted edge contains e and whose lower base passes through any point q_k with $k > j$; see Figure 8(c). If q_j is the lower-right vertex of M, the other point lying on the slanted edge belongs to $R^{(j)}$ and is uniquely determined. Hence the number of regions of type (b) (ignoring the extreme rectangular regions) is upper bounded by $r_{u,v}$ plus the overall number of distinct edges of the "incremental" convex hulls of $R^{(1)}, \ldots, R^{(r_{u,v})}$. The latter number is $O(r_{u,v})$ because every newly added point q_j generates one new edge of the modified hull, possibly deleting several other edges from the hull. (Note that this is exactly the analysis of the classical "Graham scan" convex hull algorithm.) There are only two extreme rectangular ranges of type (a) in $\mathcal{M}_{u,v}$. The number of extreme rectangular ranges of type (b) is easily seen to be $O(r_{u,v} + 1)$, using a variant of the analysis in Section 2.

Let $\mathcal{M}$ be the union of all the sets $\mathcal{M}_{u,v}$, over all primary nodes u and all nodes v of the respective secondary trees $\mathcal{T}_u$. Then we have $|\mathcal{M}| = O(|R| \log^2 r + r \log r)$.

We also have the following promised property: Let T be the remaining portion of an initial triangle T_0, and let u and v be the respective primary and secondary nodes for which T is an anchored triangle or trapezoid within $\sigma_{u,v}$, as constructed above. Then, if T is $R_{u,v}$-empty, it is contained in the union of at most two regions of $\mathcal{M}_{u,v}$.

Indeed, we may assume that $R_{u,v} \neq \emptyset$. Suppose first that T is a triangle. Translate the hypotenuse of T away from the apex o of $\sigma_{u,v}$, until it passes through a point q of $R_{u,v}$ (necessarily a hull vertex). Then rotate the new hypotenuse about q clockwise (resp., counterclockwise) until it meets a second point q' (resp., q'') of $R_{u,v}$ or becomes vertical (resp., horizontal). The two resulting triangles (or rectangles in the extreme cases) clearly belong to $\mathcal{M}_{u,v}$, and their union covers T. See Figure 9(a).

Suppose next that T is a trapezoid. Expand T downwards by sliding its bottom edge parallel to itself, while keeping the remaining bounding lines fixed, until its bottom edge hits some point $q = q_j$ of $R_{u,v}$ (that is, in the above notation, q is the jth highest point of $R_{u,v}$), or else reaches the lower boundary of $\sigma_{u,v}$. Then translate the slanted edge of the new trapezoid to the right until it hits a point q' of $R_{u,v}$ (more precisely, of $R^{(j)}$). Finally, rotate the new slanted edge about q' clockwise and counterclockwise until it meets a second point of $R^{(j)}$, or becomes vertical or horizontal; the clockwise rotation may end when it hits $q = q_j$. This yields two trapezoids (or rectangles) of $\mathcal{M}_{u,v}$ whose union covers T. See Figure 9(b). (Note that in both cases, the expansion of T may fall

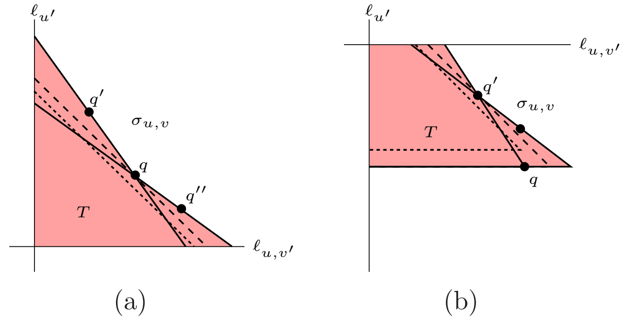


Figure 9: The dotted edges are those of the original triangle or trapezoid T. The dashed edges are the slanted edges of appropriate expansions of the original T. Each such expansion is contained in the union of a pair of regions of $\mathcal{M}_{u,v}$.

outside $\sigma_{u,v}$. This, however, does not violate our analysis, since in this case T is still contained in the union of at most two regions of $\mathcal{M}_{u,v}$, possibly clipped within $\sigma_{u,v}$.)

By construction, at least one of these two $R_{u,v}$-empty regions must contain at least n/r points of P. The analysis now continues almost verbatim as in Section 3; that is, for each heavy region $M \in \mathcal{M}_{u,v}$ with weight factor $t_M \geq c \log\log r$, we construct a $(1/t_M)$-net N_M of size $O(t_M \log t_M)$, and output the union N of R with all the resulting nets N_M. The preceding arguments, combined with the analysis in the previous sections, imply that N is indeed an ε-net. Using the Exponential Decay Lemma, which does hold in the present scenario, it can easily be shown that the expected total size of the nets N_M is sublinear in r (for the above choice of c), and thus the expected overall size of the resulting net is $O(r \log\log r)$. We have thus shown:

Theorem 4.1. *For any set P of n points in the plane, any fixed constant parameter $\alpha > 0$, and a parameter $\varepsilon > 0$, there exists an ε-net of P, for α-fat triangles, of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, where the constant of proportionality depends on α.*

Remark. Once we have restricted our attention to the case of a single semi-canonical family, the remaining analysis does not depend on any assumption concerning the slope of the hypotenuses of the triangular ranges. It thus follows that the bound in Theorem 4.1 also holds for the size of ε-nets for points in the plane and any family of triangular ranges, each of which has a pair of edges at two fixed orientations.

Constructing the ε-net. We construct an ε-net of this size using an easy variant of the algorithms presented in Sections 2 and 3. For each of the $O(1/\alpha^2)$ semi-canonical families, we apply an affine transformation to the plane (and to P), which turns the two fixed edge directions into the coordinate directions. Let us fix one such family. We construct the 2-level range tree $\mathcal{T}$ over the (transformed) points of P, using $O(n \log n)$ time and storage. We next enumerate the maximal anchored $R_{u,v}$-empty regions M in each canonical quadrant $\sigma_{u,v}$, by tracking the edges appearing on the "incremental" convex hull of the points in $R_{u,v}$, for lower-right quadrants, or by just enumerating the edges of the lower-left hull of $R_{u,v}$, for upper-right quadrants. We can produce these regions in time $O(r_{u,v} \log r_{u,v})$, although we still need to test which of them is $R_{u,v}$-empty. For simplicity, we perform this step by brute force. This takes a total of $O(1 + r_{u,v}^2)$ time per node, so the overall cost of producing the canonical empty regions is $O(s^2)$, as is easily checked (in this case the time bounds constitute a geometric sequence over the various levels of the tree); we assume s to be sufficiently small (specifically, $s = o(n^{1/2})$) to make this bound sublinear in n. Finding the degenerate canonical empty rectangles can be done by applying enumeration algorithms similar to those in Section 2.

We next compute the weight factor t_M of each of the $O(s \log^2 r)$ maximal empty regions M. For this, we prepare an appropriate version of a triangle range counting structure in the plane, which uses linear storage and $O(n \log n)$ preprocessing time, and answers queries in time $O(n^{1/2} \operatorname{polylog} n)$ [Mat92a]. The overall cost of answering $O(s \log^2 r)$ queries, including preprocessing, is $O(n \log n + sn^{1/2} \log^2 r \operatorname{polylog} n)$, which is only $O(n \log n)$, for $s = o(n^{1/2})$. We then proceed in a similar manner as that described for the previous algorithms in Sections 2 and 3. Omitting any further details, we obtain that the overall expected running time of the algorithm is $O(n \log n)$, with a constant of proportionality that depends on α (for $s = o(n^{1/2})$).

5 Improved bounds for ε-nets for other range spaces

In this section we observe that the technique developed in this paper can be adapted to the scenarios considered by Clarkson and Varadarajan [CV07], and yields improved bounds for the size of ε-nets in many of the cases considered there. As a consequence, using the same implication as in [CV07] (which is based on the technique of Brönnimann and Goodrich [BG95]), but with the improved bounds on the size of ε-nets in the respective range spaces, we obtain approximation algorithms for geometric SET COVER or HITTING SET with improved approximation factors. We list these improvements in Section 6, including similar improved approximation factors for the three primal range spaces considered so far in this paper—points and axis-parallel rectangles in the plane or boxes in 3-space, and points and α-fat triangles in the plane.

Rephrasing the notation used in the introduction, we consider the dual range space $\Xi = (\mathcal{C}, \mathcal{Q})$, where the ground set $\mathcal{C}$ is a collection of geometric regions in $\mathbb{R}^d$, and each range in $\mathcal{Q}$ is of the form $Q_x = \{C \in \mathcal{C} \mid x \in C\}$, for some $x \in \mathbb{R}^d$. Clarkson and Varadarajan [CV07] further assume that, for any finite subcollection $\mathcal{C}'$ of m regions of $\mathcal{C}$, the complement of the union of $\mathcal{C}'$ can be decomposed into at most $m\varphi(m)$ *cells* of some simple shape, where $\varphi(m)$ is some slowly increasing function; for technical reasons, we also require φ to be *sublinear*, in the sense that $\varphi(\alpha x) \le \alpha\varphi(x)$ for any integers $\alpha, x \ge 1$ (this latter property holds in all applications considered here and in [CV07]).

In addition, we assume that each cell in the decomposition is a (possibly unbounded) portion of space that is *defined* by $O(1)$ regions of $\mathcal{C}'$, in the sense that it appears in the decomposition of the complement of the union of just those $O(1)$ regions (in particular, the cells of the decomposition do not necessarily have the same shape as the regions of $\mathcal{C}$). In many geometric range spaces of this kind, the cells are those generated by the *vertical decomposition* of the complement of the union [SA95], although there exist other types of decompositions for various special classes of regions; see, e.g., [AMS98, Cla87, CS89] for a description of this (standard) setup.

Under these assumptions, Clarkson and Varadarajan show that the range space Ξ admits ε-nets of size $O\left(\frac{1}{\varepsilon}\varphi\left(\frac{1}{\varepsilon}\right)\right)$. Thus, if $\varphi(m) = o(\log m)$, the resulting nets have size smaller than the standard bound $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ of [HW87].

In this section we obtain the following improvement.[5]

Theorem 5.1. *Under the assumptions made above, the range space Ξ admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\varphi\left(\frac{1}{\varepsilon}\right)\right)$, for any $0 < \varepsilon \le 1$.*

[5] Of course, it is an improvement only when φ is $\omega(1)$; otherwise, the bound is $O(1/\varepsilon)$, as already follows from [CV07].

Remark. The bound in the theorem improves upon the general bound $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ when $\varphi(m) = 2^{o(\log m)}$, thus extending the applicability of this technique beyond the "effective range" $\varphi(m) = o(\log m)$, where the original technique of [CV07] yields an improvement.

Proof. We follow the general approach of Section 2. Here we have a finite subcollection of n elements of $\mathcal{C}$, which, for simplicity, we continue to denote by $\mathcal{C}$. We put $r := 1/\varepsilon$, $s := cr\log\varphi(r)$, and $\pi := s/n$, where $c > 1$ is a constant. We draw a random sample R of regions of $\mathcal{C}$, picking each region, independently, with probability π. We form the union $\mathcal{U}$ of R and decompose its complement into at most $|R|\varphi(|R|)$ simply-shaped regions, each determined by $O(1)$ sets of R; as above, we refer to the regions which form the decomposition as "cells". We define the weight factor t_M of a cell M to be $s|\mathcal{C}_M|/n$, where $\mathcal{C}_M$ is the subcollection of those regions of $\mathcal{C}$ which meet M. By the standard ε-net theory [HW87], or, alternatively, by the Clarkson-Shor technique [Cla87, CS89], it follows that, with high probability, we have[6] $|\mathcal{C}_M| = O\left(\frac{n}{s}\log s\right)$ for each cell M, and, in an informal and imprecise sense, the expected size of $\mathcal{C}_M$, for a cell M, is only $O(n/s)$.

As above, we take each "heavy" cell M, with $t_M \geq c\log\varphi(r)$, and use the standard theory of ε-nets to deduce that there exists a $(1/t_M)$-net N_M for $\mathcal{C}_M$, whose size is $O(t_M\log t_M)$. We output the union of R with all the sets N_M, over all heavy cells M, as the desired $(1/r)$-net (that is, ε-net) N.

Adapting the argument in Section 2, it is straightforward to verify that N is indeed an ε-net. Recall that in this dual context an ε-net is a subset of regions that cover all points that are contained in at least an ε-fraction of the regions. To bound the expected size of N, we follow the same analysis as in Section 2. That is, we apply the Exponential Decay Lemma in this context. Here, for a cell M, its defining set $D(M)$ consists of the $O(1)$ regions that determine M, and its killing set $K(M)$ is the set of all regions in $\mathcal{C}$ that intersect M. In essentially all cases considered in [CV07] and below, the axioms assumed in [AMS98], or their simplified version used in Section 2, hold. We denote by $\mathrm{CT}(R)$ the set of all cells appearing in the decomposition of the complement of the union of a subset R of $\mathcal{C}$, and by $\mathrm{CT}_t(R)$ the subset of $\mathrm{CT}(R)$ consisting of those cells with weight factor at least t.

It thus follows that the Exponential Decay Lemma is applicable in this scenario as well, and it implies that, for any t,

$$\mathbf{Exp}\left\{|\,\mathrm{CT}_t(R)|\right\} = O\left(2^{-t}\,\mathbf{Exp}\left\{|\,\mathrm{CT}(R')|\right\}\right) = O\left(2^{-t}\,\mathbf{Exp}\left\{|R'|\varphi(|R'|)\right\}\right),$$

where R (resp., R') is a random sample in which each region of $\mathcal{C}$ is chosen independently with probability s/n (resp., $s/(tn)$).

To bound the latter expectation, we argue as follows.[7] Let $z := s/t$ denote the expected value of $|R'|$. By Chernoff's bound (see, e.g., [AS92]),

$$\Pr\left\{|R'| \geq \xi z\right\} \leq e^{-(\xi-1)^2 z/3},$$

[6] Normally, for these bounds to hold, one needs to consider only those regions of $\mathcal{C}$ which *cross* (i.e., intersect but do not fully contain) M. However, in our case we do not need this distinction: Since each cell M is disjoint from all regions of R, the above analysis also applies to regions of $\mathcal{C}$ that fully contain M.

[7] Here we pay back a little for using the simpler sampling model.

for any $\xi > 1$. Hence, using the sublinearity of φ,

$$\begin{aligned}\mathbf{Exp}\left\{|R'|\varphi(|R'|)\right\} &\leq 2z\varphi(2z) + \sum_{j\geq 2} \Pr\left\{|R'| \geq jz\right\}(j+1)z\varphi((j+1)z) \\ &\leq z\varphi(z)\cdot\left(4 + \sum_{j\geq 2}(j+1)^2 e^{-(j-1)^2 z/3}\right) = O(z\varphi(z)).\end{aligned}$$

In particular, for $t = c\log\varphi(r)$, each point is chosen in R' with probability r/n (so $z = r$), and we get

$$\mathbf{Exp}\left\{|\operatorname{CT}_t(R)|\right\} = O\left(2^{-c\log\varphi(r)} r\varphi(r)\right) = O\left(\frac{r}{\varphi^{c-1}(r)}\right),$$

which, for $c > 1$, is sublinear in r. For larger values of t, the expectation is $O\left(2^{-t}(s/t)\varphi(s/t)\right)$.

We can now continue with the analysis of Section 2 almost verbatim, arguing that the overall expected size of the subsamples "within" each heavy cell of the complement of the union is sublinear in r, so the expected size of N is dominated by that of R, thus it is $O(r\log\varphi(r))$. □

Several special cases. Theorem 5.1 immediately implies improved bounds on the size of ε-nets for dual range spaces of several classes of regions and points, for which the union complexity (or, rather, the complexity of the decomposition of its complement) is known to be nearly linear. We first present some of the standard families with this property, and state their union complexity. Since these are families of planar regions, the following bounds also apply, with some care, for the complexity of the decomposition of the complement of their union. (We only consider families for which the known bound is super-linear; there is no improvement when the union complexity is linear.)

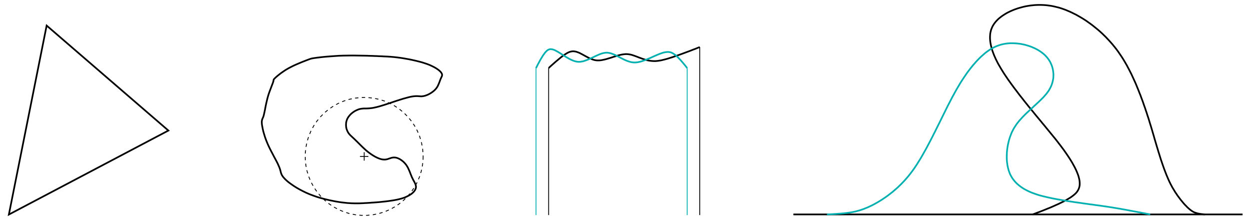
α-fat triangles (Figure 10(a)). Recall that a triangle is *α-fat* if each of its angles is at least α. The complexity of the union of n such triangles is $O(n\log\log n)$, where the constant of proportionality depends on the fatness factor α [MPSSW94, PT02].

Locally γ-fat objects (Figure 10(b)). These objects were recently introduced by de Berg [dB08]. Given a parameter $0 < \gamma \leq 1$, an object o is *locally γ-fat* if, for any disk D whose center lies in o, such that D does not fully contain o in its interior, we have $\operatorname{area}(D \sqcap o) \geq \gamma\cdot\operatorname{area}(D)$, where $D \sqcap o$ is the connected component of $D \cap o$ that contains the center of D. We also assume that the boundary of each of the given objects has only $O(1)$ locally x-extreme points, and that the boundaries of any pair of objects intersect in at most s points, for some constant s. It is then shown in [dB08] that the combinatorial complexity of the union of n such objects is $O(\lambda_{s+2}(n)\log^2 n)$, with a constant of proportionality that depends on γ. When the objects have roughly the same size (i.e., the ratio of the diameters of any pair of objects is bounded by some constant), the complexity of the union decreases to $O(\lambda_{s+2}(n))$. Locally γ-fat objects are a generalization of several other previously studied classes of "fat" objects [Ef05, EK99, ES00].

Semi-unbounded pseudo-trapezoids (Figure 10(c)). Here each object is a region of one of the forms

$$\begin{aligned}\tau^-_{x_1,x_2,f} &= \left\{(x,y) \mid x_1 \leq x \leq x_2,\ y \leq f(x)\right\}, \text{ or} \\ \tau^+_{x_1,x_2,f} &= \left\{(x,y) \mid x_1 \leq x \leq x_2,\ y \geq f(x)\right\},\end{aligned}$$

where f is a continuous function. We assume that the graphs of any pair of these functions intersect in at most s points, for some constant s. In this case the complexity of the union of any n such



(a) (b) (c) (d)

Figure 10: The types of regions considered in this section: (a) an α-fat triangle; (b) a locally γ-fat region; (c) semi-unbounded pseudo-trapezoids; and (d) regions bounded by Jordan arcs with three intersections per pair.

objects is $O(\lambda_{s+2}(n))$; see, e.g., [SA95]. If the objects are *pseudo-halfplanes*, that is, $x_1 = -\infty$ and $x_2 = +\infty$ for each object, the bound on the union complexity slightly improves to $O(\lambda_s(n))$.

Jordan arcs with three intersections per pair (Figure 10(d)). Each object is bounded by some Jordan arc which starts and ends on the x-axis but otherwise lies above it, and by the portion of the x-axis between these endpoints, and each pair of the bounding Jordan arcs intersect at most three times. In this case the complexity of the union of any n such objects is $O(\lambda_3(n)) = O(n\alpha(n))$; see [EGH*89]. We also assume that the boundary of each object has only $O(1)$ locally x-extreme points.

Recall that the actual condition is about the complexity of a decomposition of the complement of the union, rather than just the complexity of the union itself. However, since we are dealing with planar objects of the above kind, the standard vertical decomposition technique (see, e.g., [SA95]) yields a decomposition whose complexity is proportional to that of the union, so the above bounds hold for the decomposition as well.[8]

As noted by Clarkson and Varadarajan [CV07], their general approach implies that any dual range space of α-fat triangles and points admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$. Similarly, any dual range space of locally γ-fat objects and points, where the objects have roughly the same size, and each pair of object boundaries intersect in at most s points, admits an ε-net of size[9] $O\left(\lambda_s\left(\frac{1}{\varepsilon}\right)\right)$. When the objects are bounded by Jordan arcs with three intersections per pair, as defined above, the size of the net becomes $O\left(\frac{1}{\varepsilon}\alpha\left(\frac{1}{\varepsilon}\right)\right)$.

Using Theorem 5.1 we can improve each of these bounds of [CV07], and also extend the bound for the case of locally γ-fat objects of *arbitrary* sizes (a case that cannot be treated by the original technique of [CV07]). That is, we have:

Corollary 5.2. *(a) Any dual range space of α-fat triangles and points in the plane admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\log\log\frac{1}{\varepsilon}\right)$, for any $0 < \varepsilon \le 1$.*

(b) Consider a dual range space of locally γ-fat objects of arbitrary sizes in the plane and points, so that the boundary of each of the given objects has only $O(1)$ locally x-extreme points, and any pair of these boundaries meet in at most s points, for s constant. Then any such dual range space admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$, for any $0 < \varepsilon \le 1$. When these objects have roughly the same size, the bound improves to $O\left(\frac{1}{\varepsilon}\log\beta_{s+2}\left(\frac{1}{\varepsilon}\right)\right)$, where $\beta_t(1/\varepsilon) = \varepsilon\lambda_t(1/\varepsilon)$.

(c) Consider a dual range space of semi-unbounded pseudo-trapezoids and points in the plane, where, for any pair of trapezoids, the graphs of their bounding functions intersect in at most s points,

[8]This is why we need to assume that no object boundary "wiggles" too much.

[9]In fact, Clarkson and Varadarajan [CV07] applied their technique to the more restricted class of (α, β)-*covered objects* of roughly equal size (see [Ef05] for the definition and the union complexity bound), and obtained a similar bound; the same technique applies to the more general class of locally γ-fat objects.

for some constant s. Then any such dual range space admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\beta_{s+2}\left(\frac{1}{\varepsilon}\right)\right)$, for any $0 < \varepsilon \le 1$. When the pseudo-trapezoids are pseudo-halfplanes, the bound improves to $O\left(\frac{1}{\varepsilon}\log\beta_{s}\left(\frac{1}{\varepsilon}\right)\right)$.

(d) Consider a dual range space of objects and points, where each object is bounded by a Jordan arc which starts and ends on the x-axis and by the portion of the x-axis between these endpoints. Each bounding Jordan arc has only $O(1)$ locally x-extreme points, and each pair of these arcs intersect at most three times. Then any such dual range space admits an ε-net of size $O\left(\frac{1}{\varepsilon}\log\alpha\left(\frac{1}{\varepsilon}\right)\right)$, for any $0 < \varepsilon \le 1$.

Remark. Applying the known upper bounds on the quantities $\beta_s(n)$ (see [ASS89, Ni09]), we have

$$\log\beta_s(n) = \begin{cases} O\left(\alpha^{\lfloor(s-2)/2\rfloor}(n)\right), & s \ge 2 \text{ even}, \\ O\left(\alpha^{\lfloor(s-2)/2\rfloor}(n)\log\alpha(n)\right), & s \ge 3 \text{ odd}; \end{cases}$$

$\beta_1(\cdot)$ and $\beta_2(\cdot)$ are constants.

In closing, we note that, although the technique, as laid out at the beginning of this section, can be applied in principle to dual range spaces in any dimension, we have managed to apply it only to planar dual range spaces. The reason is the scarcity of classes of regions in higher dimensions with linear or, rather, near-linear bounds on their union complexity. It would of course be interesting to find such classes, and to apply to them our new technique.

6 Improved approximation factors for geometric Set Cover and Hitting Set

In this section we plug the improved bounds on the size of ε-nets, as derived in the preceding sections, into the machinery of Brönnimann and Goodrich [BG95], to obtain improved approximation factors for the corresponding SET COVER or HITTING SET problems.

We first briefly recall this technique. For simplicity, we only review the *hitting set* variant. We are given a range space $(X, \mathcal{R})$, and the goal is to find a small subset S of X which meets every range in $\mathcal{R}$. The technique assumes the availability of two black-box routines: (i) an *ε-net finder*, which is a procedure that, given any weight function w on X and $\varepsilon > 0$, constructs a *weighted* ε-net N for $(X, \mathcal{R})$, in the sense that N hits each range whose weight is at least $\varepsilon w(X)$, with the weight of a subset of X being the sum of weights of its elements; (ii) a *verifier*, which is a procedure that, given a subset $H \subseteq X$, either determines that H is a hitting set, or returns a nonempty range $R \in \mathcal{R}$ such that $R \cap H = \emptyset$.

The algorithm runs an exponential search to guess the value of OPT, which is the size of the smallest hitting set for $(X, \mathcal{R})$. At each step, denote by OPT the current guess for this value. The algorithm assigns weights (initially, uniform) to the elements of X, and uses the net finder to select a (weighted) $\frac{1}{2\text{OPT}}$-net N. If the verifier determines that N is a hitting set, it outputs N and stops. Otherwise, it returns some range R not hit by N. We double the weights of the elements in R, and repeat the above procedure. We keep iterating in this manner until N hits all the ranges in $\mathcal{R}$, or abort after a pre-specified number of iterations, concluding then (with high probability) that the current guess for OPT is too small. Thus, upon termination, the size of the reported hitting set has the same upper bound as that for $\frac{1}{2\text{OPT}}$-nets.

The analysis of [BG95] (see also [Cla93]) shows that, at the right guess for OPT, the algorithm terminates after at most $O(\text{OPT}\log(n/\text{OPT}))$ rounds. Thus the overall running time of the algorithm is (observing that since the guessed values of OPT obtained at the exponential search form a geometric sequence, the running time is dominated by that of the last round)

$$O\left(\textsc{Opt}\log\left(n/\textsc{Opt}\right)(T_N+T_V)\right),$$

where T_N, T_V are the respective running time bounds for the net-finder and the verifier.

We first apply this technique to our three main (primal) range spaces, consisting of points and axis-parallel rectangles in the plane, of points and axis-parallel boxes in $\mathbb{R}^3$, and of points and α-fat triangles in the plane. We have presented in Sections 2, 3, and 4 algorithms that construct an ε-net for these cases in nearly-linear time, and it is straightforward to generalize these algorithms to the weighted case within the same asymptotic time bound (using, e.g., the technique of Matoušek [Mat95]). The verifier can easily be implemented in polynomial time, using either brute force or some more refined range-searching machinery. We thus obtain:

Corollary 6.1. *There exists a randomized, expected polynomial-time algorithm that, given a set $\mathcal{Q}$ of m axis-parallel rectangles and set P of n points in the plane that hit $\mathcal{Q}$, computes a subset $H \subseteq P$ of $O(\textsc{Opt}\log\log\textsc{Opt})$ points that hit $\mathcal{Q}$, where $\textsc{Opt}$ is the size of the smallest such set. The algorithm can be extended to the case of axis-parallel boxes and points in 3-space, and α-fat triangles and planar point sets, yielding a similar approximation factor.*

Using the above machinery, we also obtain polynomial-time approximation algorithms for the SET COVER problems associated with the dual range spaces considered in Section 5. As shown in [CV07, Theorem 2.3], the ε-net can be constructed in time that is polynomial in the size of the ground set (and in $1/\varepsilon$), and the verifier can be implemented in polynomial time as well, either by using an appropriate range-searching machinery, or by a brute-force procedure. We thus have:

Corollary 6.2. *(a) There exists a randomized, expected polynomial-time algorithm that, given a set P of n points in the plane and a set T of α-fat triangles that cover P, computes a set cover $T' \subseteq T$ for P of size $O(\textsc{Opt}\log\log\log\textsc{Opt})$, where $\textsc{Opt}$ is the size of the smallest such set.*

(b) There exists a randomized, expected polynomial-time algorithm that, given a set P of n points in the plane and a set T of locally γ-fat objects of arbitrary sizes that cover P, so that the boundary of each of the given objects has only $O(1)$ locally x-extreme points, and each pair of these boundaries intersect in at most s points, for some constant s, computes a set cover $T' \subseteq T$ for P of size $O(\textsc{Opt}\log\log\textsc{Opt})$, where $\textsc{Opt}$ is the size of the smallest such set. When the elements of T have (roughly) the same size, the size of the set cover improves to $O\left(\textsc{Opt}\log\beta_{s+2}(\textsc{Opt})\right)$.

(c) There exists a randomized, expected polynomial-time algorithm that, given a set P of n points in the plane and a set T of semi-unbounded pseudo-trapezoids that cover P, bounded by x-monotone curves, each pair of which meet at most s times, computes a set cover $T' \subseteq T$ for P of size $O\left(\textsc{Opt}\log\beta_{s+2}(\textsc{Opt})\right)$, where $\textsc{Opt}$ is the size of the smallest such set; the bound slightly improves to $O\left(\textsc{Opt}\log\beta_{s}(\textsc{Opt})\right)$, when the input regions are pseudo-halfplanes.

(d) There exists a randomized, expected polynomial-time algorithm that, given a set P of n points in the plane and a set T of objects that cover P, each of which is bounded by some Jordan arc which starts and ends on the x-axis and by the portion of the x-axis between these endpoints, so that each bounding Jordan arc has only $O(1)$ locally x-extreme points, and each pair of these arcs intersect at most three times, computes a set cover $T' \subseteq T$ for P of size $O\left(\textsc{Opt}\log\alpha(\textsc{Opt})\right)$, where $\textsc{Opt}$ is the size of the smallest such set.

7 Concluding remarks and open problems

In this paper we achieved significant progress on the problem of bounding the size of ε-nets for several set systems, both in the primal and the dual (geometric) setting. We conclude the paper by stating several open problems raised by our study.

(i) One may consider the dual version of the main problem that we have studied. Namely, we are given a collection $\mathcal{C}$ of n axis-parallel rectangles, and each range is the subset of $\mathcal{C}$ stabbed by some point in the plane. Here too the goal is to show the existence of a small-size ε-net, which is a (small-size) subset $\mathcal{C}' \subseteq \mathcal{C}$ whose union contains all the "deep" points (i.e., points contained in at least εn rectangles of $\mathcal{C}$). So far we do not know how to apply our method to this dual setup. We note that Brönnimann and Lenchner, in their conference paper [BL04], claim, without a proof, the existence of ε-nets for this dual range space, of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$.

(ii) Another challenging open problem is to extend our machinery for axis-parallel boxes to dimensions $d \geq 4$. The anchoring trick used for $d = 3$ fails, because the number of maximal R-empty orthants in d-space can be $\Theta\left(|R|^{\lfloor d/2 \rfloor}\right)$ in the worst case [KRSV07], and the challenge is to prune away most of these orthants, and remain only with a nearly-linear number of them. A modest goal is to construct a *weak ε-net* in this setting (that is, the points in the ε-net do not necessarily have to be from the input set). This problem is addressed in a follow-up study [Ezra09], where the bound is shown to be $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$ in any dimension d (with a constant of proportionality that depends on d). Another goal is to construct weak ε-nets of size $o\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$ for the (primal) range spaces that we have studied in this paper, most notably for points and axis-parallel rectangles. In fact, it would also be interesting to find a simpler construction that yields weak ε-nets of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$.

(iii) Last but not least, there is the problem of constructing small-size ε-nets for the primal range spaces whose duals were considered in Section 5, such as those involving planar point sets and locally γ-fat objects, or semi-unbounded pseudo-trapezoids, with the properties assumed in Section 5. (We did achieve this goal for α-fat triangles.)

References

[AMS98] P. K. Agarwal, J. Matoušek, and O. Schwarzkopf, Computing many faces in arrangements of lines and segments, *SIAM J. Comput.*, 27:491–505, 1998.

[ASS89] P. Agarwal, M. Sharir, and P. Shor, Sharp upper and lower bounds for the length of general Davenport Schinzel sequences, *J. Combin. Theory, Ser. A*, 52:228–274, 1989.

[AS92] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley-Interscience, New York, 1992.

[BGLR93] M. Bellare, S. Goldwasser, G. Lund, and A. Russell, Efficient probabilistically checkable proofs and applications to approximation, In *Proc. 25th Annu. ACM Symp. Theory Comput.*, pp. 294–304, 1993.

[BEHW89] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth, Classifying learnable geometric concepts with the Vapnik-Chervonenkis dimension, *J. ACM*, 36(4):929–965, 1989.

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

[BG95] H. Brönnimann and M. T. Goodrich, Almost optimal set covers in finite VC dimensions, *Discrete Comput. Geom.*, 14:463–479, 1995.

[BL04] H. Brönnimann and J. Lenchner, Fast almost-linear-sized nets for boxes in the plane, In *Proc. 14th Annu. Fall Workshop Comput. Geom.*, pp. 36–38, 2004.

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

[Cla87] K. L. Clarkson, New applications of random sampling in computational geometry, *Discrete Comput. Geom.*, 2:195–222, 1987.

[Cla93] K. L. Clarkson, Algorithms for polytope covering and approximation, In *Proc. 3rd Workshop on Algorithms and Data Structures*, pages 246–252, 1993.

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

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

[CLRS01] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms.* MIT Press, Cambridge, MA, 2001.

[dB08] M. de Berg, Improved bounds on the union complexity of fat objects, *Discrete Comput. Geom.*, 40(1):127–140, 2008.

[dBCKO08] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd rev. ed., Springer-Verlag, Berlin, 2008.

[Ef05] A. Efrat, The complexity of the union of (α, β)-covered objects, *SIAM J. Comput.*, 34:775–787, 2005.

[EGH*89] H. Edelsbrunner, L. Guibas, J. Hershberger, J. Pach, R. Pollack, R. Seidel, M. Sharir, and J. Snoeyink, On arrangements of Jordan arcs with three intersections per pair, *Discrete Comput. Geom.*, 4(1):523–539, 1989.

[EK99] A. Efrat and M. Katz, On the union of κ-curved objects, *Comput. Geom. Theory Appl.* 14:241–254, 1999.

[Ezra09] E. Ezra, Weak ε-nets for axis-parallel boxes in d-space, *Manuscript*, 2009.

[ES00] A. Efrat and M. Sharir, The complexity of the union of fat objects in the plane, *Discrete Comput. Geom.* 23:171–189, 2000.

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

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

[FPT81] R. Fowler, M. Paterson, and S. Tanimoto, Optimal packing and covering in the plane are NP-complete, *Inform. Process. Lett.*, 12(3):133–137, 1981.

[GJ79] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, New York, NY, 1979.

[HKSS08] S. Har-Peled, H. Kaplan, M. Sharir, and S. Smorodinsky, ε-nets for halfspaces revisited, manuscript, 2008.

[HW87] D. Haussler and E. Welzl, ε-nets and simplex range queries, *Discrete Comput. Geom.*, 2:127–151, 1987.

[KRSV07] H. Kaplan, N. Rubin, M. Sharir, and E. Verbin, Efficient colored orthogonal range counting, *SIAM J. Comput.*, 38:982–1011, 2008.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher, Eds., *Complexity of Computer Computations*, pages 85–103, Plenum Press, New York, 1972.

[KPW92] J. Komlós, J. Pach, and G. Woeginger, Almost tight bounds for epsilon nets, *Discrete Comput. Geom.*, 7:163–173, 1992.

[Mat02] J. Matoušek, *Lectures on Discrete Geometry*, Springer-Verlag New York, 2002.

[Mat92a] J. Matoušek, Efficient partition trees, *Comput. Geom. Theory Appl.* 8:315–334, 1992.

[Mat92b] J. Matoušek, Reporting points in halfspaces, *Comput. Geom. Theory Appl.* 2:169–186, 1992.

[Mat95] J. Matoušek, Approximations and optimal geometric divide-and-conquer, *J. Comput Sys. Sci.*, 50(2):203–208, 1995.

[MSW90] J. Matoušek, R. Seidel, and E. Welzl, How to net a lot with little: Small ε-nets for disks and halfspaces, In *Proc. 6th Annu. ACM Sympos. Comput. Geom.*, pages 16–22, 1990. Revised version at `http://kam.mff.cuni.cz/~matousek/enets3.ps.gz` .

[MPSSW94] J. Matoušek, J. Pach, M. Sharir, S. Sifrony, and E. Welzl, Fat triangles determine linearly many holes, *SIAM J. Comput.*, 23:154–169, 1994.

[NLH84] A. Naamad, D. T. Lee, and W. L. Hsu, On the maximal empty rectangle problem, *Discrete Applied Math.*, 8:267–277, 1984.

[Ni09] G. Nivasch, Improved bounds and new techniques for Davenport-Schinzel sequences and their generalizations, *Proc. ACM-SIAM Annu. Sympos. Discrete Algorithms*, pages 1–10, 2009.

[PA95] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, Wiley Interscience, New York, 1995.

[PT02] J. Pach and G. Tardos, On the boundary complexity of the union of fat triangles, *SIAM J. Comput.*, 31:1745–1760, 2002.

[PR08] E. Pyrga and S. Ray, New existence proofs for ε-nets, *Proc. 24th Annu. ACM Sympos. Comput. Geom.*, pages 199–207, 2008.

[SA95] M. Sharir and P. K. Agarwal, *Davenport-Schinzel Sequences and Their Geometric Applications*, Cambridge University Press, New York, 1995.

[Var08] K. Varadarajan, Epsilon nets and union complexity, *Proc. 25th Annu. Sympos. Comput. Geom.*, 2009, to appear.

A Proof of the Exponential Decay Lemma

Proof of Lemma 2.1: For a fixed level i, let T denote the collection of all axis-parallel rectangles which are anchored at the entry side of some strip σ_v at that level, and each of their three other sides contains a point of P_v (or extends all the way to the strip boundary or to $\pm\infty$, as appropriate). Let T_t denote the subset of T consisting of all rectangles with weight factor at least t. We have

$$\mathbf{Exp}\left\{|\operatorname{CT}_t(R)|\right\} = \sum_{M\in T_t} \mathbf{Prob}\left\{M \in \operatorname{CT}(R)\right\}, \tag{1}$$

$$\mathbf{Exp}\left\{|\operatorname{CT}(R')|\right\} = \sum_{M\in T} \mathbf{Prob}\left\{M \in \operatorname{CT}(R')\right\} \geq \sum_{M\in T_t} \mathbf{Prob}\left\{M \in \operatorname{CT}(R')\right\}. \tag{2}$$

In view of (1) and (2), it suffices to show that, for each $M \in T_t$,

$$\mathbf{Prob}\left\{M \in \operatorname{CT}(R)\right\} = O\left(2^{-t}\right) \cdot \mathbf{Prob}\left\{M \in \operatorname{CT}(R')\right\}.$$

Let A_M be the event that $D(M) \subset R$ and $K(M) \cap R = \emptyset$, and let A'_M be the event that $D(M) \subset R'$ and $K(M) \cap R' = \emptyset$. In our setup, the event A_M is exactly the event $M \in \operatorname{CT}(R)$, and the event A'_M is exactly the event $M \in \operatorname{CT}(R')$. Moreover, putting $\delta := |D(M)| \leq 3$, $w := |K(M)|$, we have $\mathbf{Prob}\{A_M\} = \pi^\delta(1-\pi)^w$, and $\mathbf{Prob}\left\{A'_M\right\} = (\pi')^\delta(1-\pi')^w$. Hence

$$\frac{\mathbf{Prob}\left\{M \in \operatorname{CT}(R)\right\}}{\mathbf{Prob}\left\{M \in \operatorname{CT}(R')\right\}} = \frac{\mathbf{Prob}\left\{A_M\right\}}{\mathbf{Prob}\left\{A'_M\right\}} = \frac{\pi^\delta(1-\pi)^w}{(\pi')^\delta(1-\pi')^w} = t^\delta \left(\frac{1-\pi}{1-\pi'}\right)^w.$$

Substituting $\pi = s/n$, $\pi' = \pi/t$, $w \geq t \cdot n/s$, the latter expression becomes $O\left(2^{-t}\right)$, which completes the proof of the lemma. □