CS60023 Approximation and Online Algorithms	21-03-2024
ϵ-nets in Approximation Algorithms	
Lecturer: Prof. S. P. Pal	*Written by: Anirudh M*

Epsilon nets (ϵ-nets), originally introduced in computational geometry, offer a principled approach to approximating geometric structures by selecting a sparse subset of points that capture essential geometric properties of the underlying space. The core idea behind ϵ-nets lies in the creation of a representative subset such that every point in the space is either contained within the subset or is within a specified distance (ϵ) from at least one point in the subset. This concept enables the approximation of geometric properties with reduced complexity while maintaining a high degree of accuracy.

In recent years, ϵ-nets have garnered significant attention in the realm of approximation algorithms, where they have been harnessed to tackle a wide array of optimization problems spanning diverse domains such as computational biology, machine learning, network design, and beyond. By leveraging the power of ϵ-nets, approximation algorithms can efficiently navigate complex solution spaces, providing near-optimal solutions with provable guarantees on solution quality and computational efficiency.

1.1 ϵ-nets – An Introduction

Before we look at ϵ-nets, there are a few things we need to know first.

Set systems: *A pair $\Sigma = (X, R)$, where X is a set of base elements (called the ground set) and R is a collection of subsets of X (a system of sets), is called a set system. The dual set system to (X, R) is the system $\Sigma^* = (X^*, R^*)$, where $X^* = R$, and for each $x \in X$, the set $R_x := R \in R : x \in R$ belongs to R^** [GOT18]

Shattering of a Set: *A subset $A \subseteq V(H)$ is called shattered if for every $B \subseteq A$, there exists an $E \in E(H)$ such that $E \cap A = B$.* [GOT18]

Let us consider a set system (a hypergraph) $H = (X, R)$ as in the definition above. For this system, the Vapnik-Chervonenkis dimension $VC(H)$ of H is defined as the cardinality of the largest shattered subset of $V(H)$ [NY71]. Now, let for any set system (X, R) and $Y \subseteq X$, the projection of R on Y be defined as the set system $R|_Y := \{Y \cap R : R \in R\}$. Then, a crucial property of VC-dimensions is that it is hereditary – if H has VC-dimension d, then for any $Y \subseteq X$, the VC-dimension of the set system $(Y, R|_Y)$ is at most d.

[Mat02] There is one other concept we must touch upon before moving on to define ϵ-nets, just because it provides a better intuition of what it is all about. The *transversal* of a hypergraph or a set system H is a set $T \subseteq X$ such that any set $F \in R$ has to intersect T, i.e, $T \cap F \neq \Phi$ for any $F \in R$. Thus, we can view the transversal of a hypergraph H as the hitting set of the set system forming H.

Now, we can easily intuit that a large hitting set is easier to find than a small hitting set, because it is pretty obvious that the larger the hitting set, the more elements it has and the probability that any given set has at least 1 element which is also in the hitting set increases tremendously. Conversely, it is easier to hit a large set with a given hitting set than a small one, because again the probability that any element in the large set is also in the hitting set is much higher than for a small set. Hence, using the same intuition, we define the concept of an ϵ-net on a finite set X as follows:

Definition 1.1. *[HW87; Mat02] Let (X, R) be a set system with X finite and let $\epsilon \in [0, 1]$ be a real number. A set $N \subseteq X$ (not necessarily one of the sets of R) is called an ϵ-net for (X, R) if $N \cap S \neq \Phi$ for all $S \in R$ with $|S| \geq \epsilon|X|$.*

So it is that an ϵ-net is a transversal (or a hitting set) for all sets larger than $\epsilon|X|$. It is also convenient to write $\frac{1}{r}$ instead of ϵ, so they may also be called $\frac{1}{r}$-nets too, where $r > 1$ is a real, positive parameter.

To extend the definition to an infinite set X, we need to use probabilistic measures to define the net rather than a size of a set, as counting the size of an infinite set is impossible. So, let us consider an arbitrary probability measure μ on the ground set X. This can take the form of a uniform distribution, or as a suitable multiple of the Lebesgue measure restricted to some geometric figure. [Mat02]

Definition 1.2. *Let X be a set, let μ be a probability measure on X, let R be a system of μ-measurable subsets of X, and let $\epsilon \in [0, 1]$ be a real number. A subset $N \subseteq X$ is called an ϵ-net for (X, R) with respect to μ if $N \cap S \neq \Phi$ for all $S \in R$ with $\mu(S) \geq \epsilon$.*

In a geometric sense, the concept of ϵ-nets is treated quite differently. An ϵ-net can be thought of as a subset of points in a metric space that efficiently represents the entire space within a certain distance ϵ, i.e, every point in the entire defined metric space is within an ϵ distance of some point within the ϵ-net. So, it provides a very good way to produce computationally efficient approximations to problems regarding complex geometric optimizations.

In a probabilistic sense, if P is a probability distribution over a set X of events, an ϵ-net for a class of subsets of X ($H \subseteq 2^{|X|}$) is some subset $S \subseteq X$ such that for any $h \in H$,

$$P(h) \geq \epsilon \implies S \cap h \neq \Phi$$

Before we move on further, we have to look at some seminal results that made work on such ϵ-nets possible.

1.2 Sauer's Lemma – The Shatter Function Lemma

Theorem 1.1. *For any set system F of VC-dimension at most d, we have $\pi_F(m) \leq \Phi_d(m)$ for all m, where $\Phi_d(m) = \sum_{i=0}^{d} \binom{m}{i}$. Here, $\pi_F(m)$ is the maximum number of distinct intersections of the sets of F with an m-point subset of X.*

1.2.1 Proof 1 [Mat02]

Let us take a general hypergraph $H = (X, R)$. Since the VC-dimension of H does not increase on passing to a subsystem of the hypergraph H, it suffices to show that any hypergraph of VC-dimension at most d on an n-point set has no more than $\Phi_d(n)$ sets. We proceed by induction on d, and for a fixed d we use induction on n.

Let H be such that $|X| = n$ and $VCdim(R) = d$, and let us fix some element $x \in X$. In the inductive step, we remove x from the ground set and work on the set system $R_1 = R|_{X \setminus \{x\}}$ on $n-1$ points. This R_1 has a VC-dimension at most d, as a sub-hypergraph can not have a larger shattering set than its super-hypergraph, and hence $|R_1| < \Phi_d(n-1)$ by the inductive hypothesis (that any set system with exactly d as it's VC-dimension and less than n vertices satisfies the above property). To see how many sets are lost when transitioning from R to R_1 in such a fashion, we notice that the only way that the number of sets decreases by removing x from X is when two sets $S_1, S_2 \in R$ give rise to the same set in R_1, which means that $S_2 = S_1 \cup \{x\}, \quad x \notin S_1$, or the other way round. This suggests that we define an auxiliary set system R_2 consisting of all sets in R_1 that correspond to such pairs $S_1, S_2 \in R$. Let this be defined as $R_2 = \{S \in R : x \notin S, S \cup \{x\} \in R\}$.

By the above discussion, we have $|R| = |R_1| + |R_2|$. Crucially, we observe that $VCdim(R_2) \leq d - 1$, since if $A \subseteq X \setminus x$ is shattered by R_2 , then $A \cup x$ is shattered by R. Therefore, $|R_2| \leq \Phi_{d-1}(n-1)$. Thus, we arrive at the recurrence

$$|R| \leq \Phi_d(n-1) + \Phi_{d-1}(n-1)$$

This can easily be proven to be $\Phi_d(n-1) + \Phi_{d-1}(n-1) = \Phi_d(n)$, and thus $|R| \leq \Phi_d(n)$. $\square$

Some points to note here are that the shatter function $\pi_F(m)$ is closely related to the edges of the hypergraph defined by (X, R), as $\pi_F(m) \leq |R|$. This is because each edge in a hypergraph is a set of points that can be shattered in the equivalent set system defined by it. So, for any $Y \subseteq X$ such that $|Y| = m$, since the set $R|_Y$ has to be a subset of R itself, $\pi_F(m)$ which is the largest number of subsets in $R|_Y$ has to be less than R, thus proving the claim.

Another point to note is that for the n-vertex hypergraph with $E(H)$ hyperedges and VC-dimension d, the bound turns out to be $|E(H)| \leq \sum_{i=0}^{d} \binom{n}{i}$ and it cannot be improved upon.

1.2.2 Proof 2 [PA02]

(This is for the case where $n = |X|$ and $\pi_F(n) = E(H)$)

Let $E(H) = \{E_i | 1 < i < m\}$, and let $X_j, 1 \leq j \leq \sum_{i=0}^{d} \binom{n}{i}$ be a list of all subsets of $V(H)$ of size at most d. Define an $m \times \sum_{i=0}^{d} \binom{n}{i}$ matrix A by:

$$a_{ij} = \begin{cases} 1 & \text{if} X_j \subseteq E_i \\ 0 & \text{if} X_j \nsubseteq E_i \end{cases}$$

Suppose, for contradiction, that $m > \sum_{i=0}^{d} \binom{n}{i}$. Then the rows of A are linearly dependent over the reals; thus there exists a nonzero function $f : E(H) \to R$ such that

$$\sum_{A \subseteq E_i} f(E_i) = \alpha \neq 0$$

. It can be seen that sets A with nonzero sums certainly exist, for we get a nonzero sum for any maximal element A of the family $\{A \in E(H) | f(A) \neq 0\}$. Obviously, $|A| \geq d+1$. Given any $A \subseteq B$, let

$$F(B) = \sum_{E_i \cap A = B} f(E_i)$$

Thus, $F(A) = \alpha$ and setting $B = A - \{a\}$ for any fixed $a \in A$

$$F(B) = \sum_{B \subseteq E_i} f(E_i) - \sum_{A \subseteq E_i} f(E_i) = -\alpha$$

and in general, if B is any ($|A|$ - k)-element subset of A, $F(B) = (-1)^k \alpha \neq 0$. This yields, in particular, that there exists at least one hyperedge E_i with $E_i \cap A = B$. Thus, A is shattered, contradicting our assumption that $VCdim(H) = d$. □

Note: the shatter function for any set system is either 2^m for all m (the case of infinite VC-dimensions) or it is bounded by a fixed polynomial.

1.3 The ϵ-net theorem

Theorem 1.2. *If X is a set with a probability measure μ, F is a system of μ-measurable subsets of X with $VCdim(F) \leq d$, $d \geq 2$, and $r \geq 2$ is a parameter, then there exists a $\frac{1}{r}$-net for (X, F) with respect to μ of size at most $Cdrlnr$, where C is an absolute constant.*

1.3.1 Proof [Mat02]

Let us put $s = Cdr\ln r$ (assuming without harm that it is an integer), and let N be a random sample picked by s independent random draws, where each element is drawn from X according to the probability distribution μ. (So the same element can be drawn several times; this does not really matter much, and this way of random sampling is chosen to make calculations simpler.) The goal is to show that N is a $\frac{1}{r}$-net with a positive probability.

To simplify formulations, let us assume that all $S \in F$ satisfy $\mu(S) \geq \frac{1}{r}$; this is no loss of generality, since the smaller sets do not play any role. The probability that the random sample N misses any given set $S \in F$ is at most $(1 - \frac{1}{r})^s \leq e^{-\frac{s}{r}}$, and so if s were at least $rln(|F| + 1)$, say, the conclusion would follow immediately. But r is typically much smaller than $|F|$ (it can be a constant, say), and so we need to do something more sophisticated.

Let E_0 be the event that the random sample N fails to be a $\frac{1}{r}$-net, i.e., misses some $S \in F$. We bound Prob$[E_0]$ from above using the following thought experiment.
By s more independent random draws we pick another random sample M. We put $k = \frac{s}{2r}$ again assuming that it is an integer, and we let E_1 be the following event:

There exists an $S \in F$ with $N \cap S = \Phi$ and $|M \cap S| \geq k$. In this event, the random sample M intersects some $S \in F$ at at least k points, while N completely misses S.

Here an explanation concerning repeated elements is needed. Formally, we regard N and M as sequences of elements of X, with possible repetitions, so $N = (X_l, X_2, ...X_s)$,

$M = (Y_l, Y_2, ...Y_s)$. The notation $|M \cap S|$ then really means $|\{i \in 1, 2, ..., s : y_i \in S\}|$, and so an element repeated in M and lying in S is counted the appropriate number of times.

Clearly, $\text{Prob}[E_1] \leq \text{Prob}[E_0]$, since E_1 requires E_0 plus something more. We are going to show that $\text{Prob}[E_1] \geq \frac{1}{2}\,\text{Prob}[E_0]$. Let us investigate the conditional probability $\text{Prob}[E_1|N]$, that is, the probability of E_1 when N is fixed and M is random. If N is a $\frac{1}{r}$-net, then E_1 cannot occur, and $\text{Prob}[E_0|N] = \text{Prob}[E_1|N] = 0$.

So, suppose that there exists an $S \in F$ with $N \cap S = \Phi$. There may be many such S, but let us fix one of them and denote it by S_N. We have $\text{Prob}[E_1|N] \geq \text{Prob}[|M \cap S_N| \geq k]$, because the second probability is a fixed case of the first. Since $\text{Prob}[|M \cap S_N| \geq k] \geq \frac{1}{2}$ by virtue of the binomial distribution Chernoff-type tail estimate of the variable $|M \cap S_N|$, $\text{Prob}[E_0|N] \leq 2\text{Prob}[E_1|N]$ for all N, and thus $\text{Prob}[E_0] \leq 2\text{Prob}[E_1]$.

Next, we are going to bound $\text{Prob}[E_1]$ differently. Instead of choosing N and M at random directly as above, we first make a sequence $A = (z_1, z_2, ...z_{2s})$ of $2s$ independent random draws from X. Then, in the second stage, we randomly choose s positions in A and put the elements at these positions into N, and the remaining elements into M (so there are $\binom{2s}{s}$ possibilities for A fixed). The resulting distribution of N and M is the same as above. We now prove that for every fixed A, the conditional probability $\text{Prob}[E_1|N]$ is small. This implies that $\text{Prob}[E_1]$ is small, and therefore $\text{Prob}[E_0]$ is small as well.

So let A be fixed. First let $S \in F$ be a fixed set and consider the conditional probability $P_S = \text{Prob}[N \cap S = \Phi, |M \cap S| \geq k|A]$. This is the probability that given such a sampling A, the samples N and M made by the strategy outlined above would yield the event E_1. If $|A \cap S| < k$, then $P_S = 0$ because you can't have an M made from A that can intersect at least k times with S then. Otherwise, we bound $P_S \leq Prob[N \cap S = \Phi|A]$. The latter is the probability that a random sample of s positions out of $2s$ in A avoids the at least k positions occupied by elements of S. This is at most

$$\frac{\binom{2s-k}{s}}{\binom{2s}{s}} \leq (1 - \frac{k}{2s})^s \leq e^{-(k/2s)s} = e^{-k/2} = e^{-Cdlnr/4} = r^{-Cd/4}$$

This was an estimate of P_S for a fixed $S \in F$. Now, finally, we use the assumption about the VC-dimension of F, via the shatter function lemma: The sets of F have at most $\Phi_d(2s)$ distinct intersections with A. Since the event $N \cap S = \Phi$ and $|M \cap S| \geq k$ depends only on $A \cap S$, it suffices to consider at most $\Phi_d(2s)$ distinct sets S, and so for every fixed A,

$$Prob[E_1|A] \leq \Phi_d(2s) \times r^{-Cd/4} \leq (\frac{2es}{d})^d r^{-Cd/4} = (2erlnr.r^{-C/4})^d < \frac{1}{2}$$

if $d, r \geq 2$ and C is sufficiently large. So $\text{Prob}[E_0] \leq 2\text{Prob}[E_1] < 1$, thus proving the theorem. □

A corollary of the above theorem exists, by virtue of the fact that such a $\frac{1}{r}$-net constructed has to be a transversal of the set system F (or R) considered. Hence,

Corollary 1.2.1. *Let R be a finite set system on a ground set X with $VCdim(R) \leq d$. Then we have $\tau(F) \leq Cd\tau^*(F)ln(\tau^*(F))$, where C is a constant as in the ϵ-net theorem.*

Here, $\tau(F)$ is the transversal number of the hypergraph (X, R) and $\tau^*(F)$ is its fractional transversal number.

1.4 ϵ-nets for Geometric Set Systems

The set systems induced by geometric objects in the d-dimensional metric space $\mathbb{R}^d$ are termed as geometric set systems. Indeed, for both primal and dual set systems induced in such a manner, the existence of ϵ-nets of size $O(\frac{1}{\epsilon}log(\frac{1}{\epsilon}))$ follows from the breakthroughs of Clarkson and Shor [CS89]. In these cases, all known asymptotic bounds on sizes of ϵ-nets follow from Theorem 1.3 (given below) and bounds on shallow-cell complexity. The relevance of shallow-cell complexity for ϵ-nets was realized after considerable effort was spent on inventing a variety of specialized techniques for constructing ϵ-nets for geometric set systems. These techniques and ideas have their own advantages, often yielding algorithms with low running times and low constants hidden in the asymptotic notation.

Theorem 1.3. *[Var10] Let (X, R) be a set system with shallow-cell complexity $\phi_R(.)$, where $\phi_R(n) = O(n^d)$ for some constant d. Let $\epsilon > 0$ be a given parameter. Then there exists an ϵ-net for R of size $O(\frac{1}{\epsilon}log(\phi_R(\frac{1}{\epsilon})))$. Furthermore, such an ϵ-net can be computed in deterministic polynomial time.*

1.5 Applications of ϵ-nets

1.5.1 Spanning Trees with Low Stabbing Number [PA02]

Let S be a set of n points in the plane. A tree T whose vertices are the points of S and whose edges are straight-line segments is called a spanning tree of S. The stabbing number of a spanning tree T, denoted by $\sigma(T)$, is defined as the maximum number of segments in T that can be intersected by a single line. Welzl gave a beautiful theorem, improved upon by Chazelle, that computes efficiently the spanning tree with a low stabbing number for a set of points in a plane.

Theorem 1.4. *(Welzl) For any set S of n points in $\mathbb{R}^2$ in general position, there is a spanning tree T whose stabbing number is at most $c\sqrt{n}log(n)$, where c is an appropriate constant.*

The proof of the same, which can be found in [PA02] pages $258 - 259$ is not given below. However, a notable lemma is used in the proof that bears mention. This lemma states that:

Theorem 1.5. *Given a set S of $n > 2$ points in $\mathbb{R}^2$ in general position, and a multiset L of m lines avoiding these points, one can always find a pair $x, y \in S$ such that the segment connecting x and y intersects at most $c'(m/\sqrt{n})log(n)$ lines of L, where c' is a suitable constant.*

It's proof is as follows:

One can assume without loss of generality that L is a set (i.e., it contains no repeated elements); otherwise, one can slightly change the position of some lines to make it so. Let us consider a subspace of a range space restricted to L (d = 2). That is, let $\Sigma_L = (L, \{R_s \cap L|$ for all open segments $s\})$, where $R_s \cap L$ is the set of all lines of L intersected by a given segment s. Now, Σ_L has finite VC-dimension d. Hence, by Corollary 1.2.1, there exists a constant $c > 1$ such that Σ_L has an ϵ-net of size at most $c(\frac{1}{\epsilon})log(\frac{1}{\epsilon})$ for any $0 < \epsilon < 1$.

Setting $\epsilon = c(logn/\sqrt(n))$, we find that there exists a subset $L' \subseteq L$ with

$$|L'| \leq c(\frac{1}{\epsilon})log(\frac{1}{\epsilon}) \leq \frac{\sqrt{n}}{2}$$

and every open segment crossing at least $\epsilon m = c(m/\sqrt{n})log(n)$ elements of L intersects at least one line belonging to L'. The lines of L' divide the plane into fewer than n cells. By the pigeonhole principle, at least one of them contains two points $x, y \in S$. The segment connecting them avoids every line of L' therefore, it cannot intersect more than $c(m/\sqrt{n})log(n)$ elements of L. $\square$

This has been used in the construction of a data structure for the following application.
Let S be a fixed set of n points in $\mathbb{R}^2$ in general position. We wish to build a linear-size data structure that allows us to answer efficiently questions (queries) of the following type: Given an open half-plane h , what is the number of points of S lying in h?

Without proof, the result is summarized here as follows: (the proof of the same can be found at [PA02], pages 260-261).

Corollary 1.5.1. *Given a set S of n points in the plane in general position, there exists a data structure using $O(nlog(n))$ space that allows us to compute $|S \cap h|$ for any query half-plane h in time $O(\sqrt{n}log^3(n)$.*

Another result that can be easily established using the concepts of ϵ-nets as discussed above is as follows:

Corollary 1.5.2. *Given a set L of n lines in general position and an integer $r > 3$, the plane can be decomposed into at most r^2 triangles (some of which are unbounded) such that the interior of each triangle intersects at most $c(n/r)logr$ lines of L, where c is a suitable constant.*

1.5.2 Spatial Partitioning [GOT18]

Consider the set system (H, R) where the base set H is a set of n hyperplanes in $\mathbb{R}^d$, and R is the set system induced by intersection of simplices in $\mathbb{R}^d$ with H. An ϵ-net for R consists of a subset H' such that any simplex intersecting at least ϵn hyperplanes of H intersects a hyperplane in H'. This implies that for any simplex Δ lying in the interior of a cell in the arrangement of H', the number of hyperplanes of H intersecting Δ is less than ϵn. One can further partition each cell in the arrangement of H' into simplices, leading to the powerful concept of cuttings. After a series of papers in the 1980s and early 1990s [CF90; Mat91], the following is the best result in terms of both combinatorial and algorithmic bounds.

Theorem 1.6. *[Cha93] Let H be a set of n hyperplanes in $\mathbb{R}^d$, and $r \geq 1$ a given parameter. Then there exists a partition of $\mathbb{R}^d$ into $O(r^d)$ interior-disjoint simplices, such that the interior of each simplex intersects at most $\frac{n}{r}$ hyperplanes of H. These simplices, together with the list of hyperplanes intersecting the interior of each simplex, can be found deterministically in time $O(nr^{d-1})$.*

Another well-known result in this same regard is Clarkson's method of constructing a spatial partitioning data structure to answer half-plane queries [Cla88]. Without proof, we state the notable results as follows:

- The initial observation for the data structure is that if we want to find a closest site in S to a point p, then knowing a closest site to p in some $X \subseteq S$ can help restrict the search in S.
- For a query point p and region A with $p \in A$, the set $S \cap \bar{C}(A)$ contains all the closest sites to p. If q is a closest site in X to p, then the candidate set $S \cap C(A)$ contains all sites closer to p than q. The key idea is to find some $X \subseteq S$, and some collection of regions, such that for every region A in the collection, the candidate set of A relative to X contains few sites.
- Such a collection of regions can be found using random sampling, as follows: take a random sample X of the sites, determine the Voronoi diagram $V(X)$ of that sample, and then compute $\Delta(V(X))$, a triangulation of the Voronoi diagram.
- The resulting set is a collection of simple regions satisfying the properties:
- The union of the regions covers E^d, the d-dimensional Euclidean space.
- The number of regions is $O(r^{[d/2]})$, for $r \to \infty$, where r is the size of X.
- With high probability, the candidate sets $S \cap C(A)$ are "small" for all regions $A \in \Delta(V(X))$, specifically, $|S \cap C(A)| - nO(log(r)/r)$ as $r \to \infty$;
- The regions in $\Delta(V(X))$ are simple, so that for point p and $A \in \Delta(V(X))$, we can tell in $O(1)$ time if $p \in A$, for fixed dimension d.
- For each $A \in \Delta(V(X))$, there is a site $q \in X$ such that all points in A are as close to q as to any other site in X.
- These properties suggest a two-step process for answering closest-point queries: given query point p, determine a region $A \in \Delta(V(X))$ that contains it, then determine the closest site to p in $R \cup (S \cap C(A))$ by linear search. For a suitable sample size, with high probability this procedure is faster than directly searching S. By repeatedly taking random samples until a sample is found for which the corresponding candidate sets are all small, a data structure with an improved worst-case query time can be constructed. Since a random sample will satisfy this condition with high probability, on average only $O(1)$ sampling repetitions need be done. Rather than search the candidate sets in linear time, this construction can be applied recursively, using a sample size r that is independent of the number

of sites. The resulting search structure is an RPO tree, in which the number of children of a node is independent of the number of sites, as is the size of the set of sites associated with each leaf node.

1.5.3 Derandomization [GOT18]

A natural use of ϵ-nets is derandomization, since by their construction they are a sampling of a space that efficiently represents the space within a certain distance of ϵ. Indeed, as we can see in [Cha96], a very efficient linear-time deterministic algorithm can be developed for Clarkson's randomized spatial partitioning algorithm. The result, again stated without proof, is as thus:

Theorem 1.7. *Let (H, w) be an LP-type problem with n constraints of combinatorial dimension at most D (the VC-dimension of the set system), satisfying Computational assumptions* 1 *and* 2*, with $\tilde{D} \leq D$ (else replace D by $max\{D, \tilde{D}\}$ below). Then the optimum of (H, w) can be found deterministically in time at most $C(D)n$, where $C(D) = O(D)^{7D}log^D(D)$.*

where the computational assumptions 1 and 2 are stated below:

Assumption 1: (Violation test) Given a basis B and a constraint $h \in H$, decide whether h violates B and return an error message if the input set B is not a basis for H.

Assumption 2: (Oracle test) A subsystem oracle for (H, R) of dimension $\tilde{\mathrm{D}}$ is available.

Now, as ϵ-nets can be constructed very efficiently in linear, deterministic time, the above outlined derandomiizaton algorithm also works very efficiently in terms of the net construction and further basis checking.

1.5.4 Hitting Sets and Set Covers

By the definition of an ϵ-net, it is a hitting set for a range space (a hypergraph) defined by (X, R). So, a natural application of ϵ-nets is to solve the hitting set problem for set systems of finite VC-Dimensions. Indeed, by the result of Brönniman and Goodrich [BG95], we can say that if a set system allows for an ϵ-net of size $rf(r)$, where $r = \frac{1}{\epsilon}$, then the hitting set problem can be solved with an approximation guarantee of $f(OPT)$, with the same function f as used in the size of the ϵ-net. The result, stated without proof, states that:

Theorem 1.8. *Let (X, R) be a set system given by a witness subsystem oracle of degree D, which admits a hitting set of size c. Let d stand for the VC-exponent (VC-dimension) of (X, R), and assume that Φ is not in R. Then a hitting set for (X, R) of size $O(dclog(dc))$ can be found in $O(nc^{D+1}log^D(dc)log(n/c))$ time in the log-RAM model.*

Since the hitting set returned is of size $O(dclog(dc))$, and the ϵ-net constructed in its computation is of size $drlog(dr)$, it is easy to see that the hitting set returned is of size $k * cf(c)$, if the net is of size $rf(r)$. Hence, the approximation guarantee provided is $f(c) = f(OPT)$.

Similarly, using the VC-dimension and the witness subsystem oracle of the dual set system rather than the set system itself, we can get the same bound on the size of the set cover returned for the SET COVER problem. This works so because the dual formulation of the HITTING SET problem is the SET COVER, and so any formulation for the HITTING SET problem if applied on the dual of its set system will generate a set cover for the set system. This is also outlined in [BG95].

1.5.5 Convex body approximations with Polytopes[SM12]

Approximating convex bodies by polytopes is a fundamental problem in computational geometry literature. The problem is to determine the minimum number of vertices n (alternatively, the minimum number of facets) of an approximating polytope for a given error $\epsilon > 0$. Error is commonly measured through the Hausdorff distance. Given a convex body K in d-dimensional space, we say that a polytope P ϵ-approximates K if the Hausdorff distance between K and P is at most ϵ.

Given a convex polytope P, defined as the intersection of halfspaces, an ϵ-approximate polytope membership query determines whether a query point q lies inside or outside P, but it may return either answer if q's distance from P's boundary is at most ϵ. Polytope membership queries find applications in many geometric areas, such as linear programming queries, ray shooting, nearest neighbor searching, and the computation of convex hulls.

Without proof, we state here that such ϵ-approximate queries can be answered via the construction of a data structure using ϵ-nets.

Theorem 1.9. *Let $K \subset \mathbb{R}^d$ be a convex body such that the width of K in any direction is at least ϵ. Then, there exists an ϵ-approximating polytope P with number of facets $O(rlogr)$, for $r = \frac{\sqrt{area(K)}}{\epsilon^{(d-1)/2}}$.*

Such a polytope P can be constructed using ϵ-nets, since the range spaces of all the ϵ-dual caps of such a body K have a finite, constant VC-dimension.

1.6 ϵ-approximations – A Tighter version of ϵ-nets [GOT18]

Definition 1.3. *Given a finite set system (X, F), and a parameter $0 \geq \epsilon \geq 1$, a set $A \subseteq X$ is called an ϵ-approximation if, for each $R \in F$,*

$$|\frac{|R|}{|X|} - \frac{|R \cap A|}{|A|}| \leq \epsilon$$

A basic theorem, given by B. Chazelle in The Discrepancy Method, states that

Given a finite set system (X, F) and a parameter $0 < \epsilon \leq 1$, an ϵ-approximation for (X, F) of size $O(\frac{1}{\epsilon^2} log(|F|))$ can be found in deterministic $O(|F| \times |X|)$ time.

One of the main uses of ϵ-approximations is in constructing a small-sized representation or "sketch" A of a potentially large set of elements X with respect to an underlying set system F. Then data queries from F on X can instead be performed on A to get provably approximate answers. For example, let us aim to pre-process a finite set X of points in the plane, so that given a query half-space h, we can efficiently return an approximation to $|h \cap X|$. For this data structure, one could use an ϵ-approximation $A \subseteq X$ for the set system (X, F) induced by the set of all half-spaces in $\mathbb{R}^2$. Then given a query half-space h, simply return $|h \cap A| \times |X|$; this answer differs from $|h \cap X|$ by at most $\epsilon \times |X|$. This enables the use of ϵ-approximations for computing certain estimators on geometric data; e.g., a combinatorial median $q \in \mathbb{R}^d$ for a point set X can be approximated by the one for an ϵ-approximation, which can then be computed in near-linear time. Indeed, this was shown by Matoušek in "Computing the center of planar point sets", and the result is stated without proof here as follows.

Theorem 1.10. *Let $X \subset \mathbb{R}^d$ be a finite point set, $0 \leq \epsilon < 1$ be a given parameter, and A be an ϵ-approximation for the primal set system induced by half-spaces in $\mathbb{R}^d$ on X. Then any center-point (median) for A is an ϵ-center-point for X.*

Another application of ϵ-approximations is in solving the shape-fitting problem.

Definition 1.4. *Let the shape family F contain, as its elements, all possible k-point subsets of $\mathbb{R}^d$; that is, each $R \in F$ is a subset of $\mathbb{R}^d$ consisting of k points. If the function $dist(.,.)$ is the Euclidean distance, then the corresponding shape fitting problem $(\mathbb{R}^d, F, dist)$ is the well-known k-median problem. If $dist(.,.)$ is the square of the Euclidean distance, then the shape fitting problem is the k-means problem. If the shape family F contains as its elements all hyperplanes in $\mathbb{R}^d$, and $dist(.,.)$ is the Euclidean distance, then the corresponding shape fitting problem asks for a hyperplane that minimizes the sum of the Euclidean distances from points in the given instance $P \subset \mathbb{R}^d$.*

Given an instance P, and a parameter $0 < \epsilon < 1$, an ϵ-coreset (S, w) "approximates" P with respect to every shape R in the given family F. Such an ϵ-coreset can be used to find a shape that approximately minimizes $\sum_{p \in P} dist(p, R)$: one simply finds a shape that minimizes $\sum_{q \in S} w(q) dist(q, R)$. For this approach to be useful, the size of the coreset needs to be small as well as efficiently computable. Building on a long sequence of works, Feldman and Langberg (see also Langberg and Schulman) showed the existence of a function $f : \mathbb{R} \to \mathbb{R}$ such that an ϵ-approximation for a carefully constructed set system associated with the shape fitting problem $(\mathbb{R}^d, F, dist)$ and instance P yields an $f(\epsilon)$-coreset for the instance P. For many shape fitting problems, this method often yields coresets with size guarantees that are not too much worse than bounds via more specialized arguments.

References

[NY71] Vapnik V. N. and Chervonenkis A. Ya. "On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities". In: *Theory of Probability and Its Applications* 16 (1971).

[HW87] D. Haussler and E. Welzl. "Epsilon-nets and simplex range queries". In: *Discrete Computational Geometry* 2 (1987), pp. 127–151.

[Cla88] Kenneth L. Clarkson. “Cutting hyperplane arrangements”. In: *SIAM Journal on Computing* 17 (1988), pp. 830–847.

[CS89] K.L. Clarkson and P.W. Shor. “Application of random sampling in computational geometry”. In: *Discrete Computational Geometry* 4 (1989), pp. 387–421.

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

[Mat91] J. Matoušek. “Cutting hyperplane arrangements”. In: *Discrete Computational Geometry* 6 (1991), pp. 385–406.

[Cha93] B. Chazelle. “Cutting hyperplanes for divide-and-conquer”. In: *Discrete Computational Geometry* 9 (1993), pp. 145–158.

[BG95] H. Brönnimann and M. T. Goodrich. “Almost optimal set covers in finite VC-dimension”. In: *Discrete & Computational Geometry* 14 (1995), pp. 463–479.

[Cha96] B. Chazelle. “On linear-time deterministic algorithms for optimization problems in fixed dimension”. In: *Journal of Algorithms* 21 (1996), pp. 579–597.

[Mat02] Jiří Matoušek. *Lectures on Discrete Geometry.* Springer Science & Business Media, 2002.

[PA02] Janos Pach and Pankaj K. Agarwal. *Combinatorial Geometry.* Springer Science & Business Media, 2002.

[Var10] K. Varadarajan. “Weighted geometric set cover via quasi uniform sampling”. In: *Proceedings of the forty-second ACM symposium on Theory of computing* (2010), pp. 641–648.

[SM12] G.D. da Fonseca S. Arya and D.M. Mount. “Polytope approximation and the Mahler volume”. In: *Proceedings of the twenty-third ACM-SIAM symposium on Discrete Algorithms* (2012), pp. 29–42.

[GOT18] Jacob E. Goodman, Joseph O’Rourke, and Csaba D Tóth. *Handbook of Discrete and Computational Geometry.* 3rd ed. CRC Press, 2018.