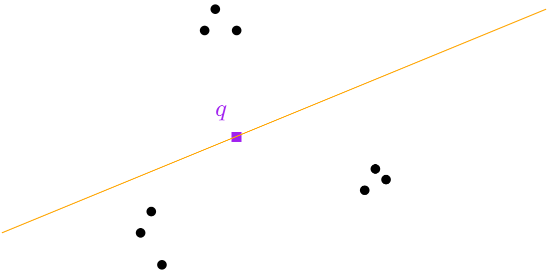
Small weak epsilon-nets

Vladyslav Rachek

March 26,2020

Introduction

Let P be a set of n points in $\mathbb{R}^2$. A point q (not necessarily in P) is called a **centerpoint** of P if each closed half-plane containing q contains at least $\lceil \frac{n}{3} \rceil$ points of P, or, equivalently, any convex set that contains more than $\frac{2}{3}n$ points of P must also contain q.



Convex sets and centerpoint

Theorem (Helly's theorem)

For any $d \geq 1, n \geq d+1$ and any family of convex sets $C_1, \ldots, C_n$ in $\mathbb{R}^d$, if intersection of any $d+1$ of these sets is non-empty, then all sets C_i intersect.

Definition. Let X be an n-point set in $\mathbb{R}^d$. A point x is called a *centerpoint* of X if each closed half-space containing x contains at least $\frac{1}{d+1}$ points of X

Convex sets and centerpoint

Theorem (Centerpoint theorem)

Each finite point set in $\mathbb{R}^d$ *has at least one centerpoint*

Proof.

We first note the *equivalent definition of a centerpoint*: x is a centerpoint of $X \subset \mathbb{R}^d$ iff it lies in each open half-space γ s.t. $|\gamma \cap X| > \frac{d}{d+1}|X|$
How do we prove that any set X has a centerpoint?
We would like to apply Helly's theorem, to conclude that all such half-spaces intersect. But we can't, because there are infinitely many such γ's, and they are open and unbounded.

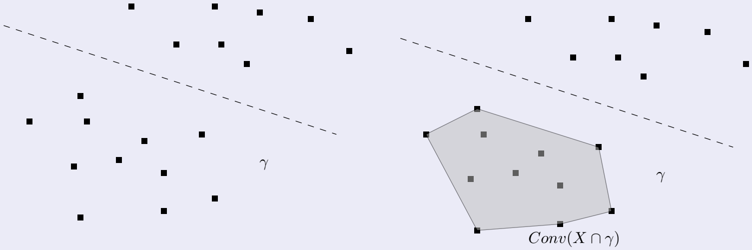
Convex sets and centerpoint

Theorem (Centerpoint theorem)

Each finite point set in $\mathbb{R}^d$ *has at least one centerpoint*

Continuation of proof.

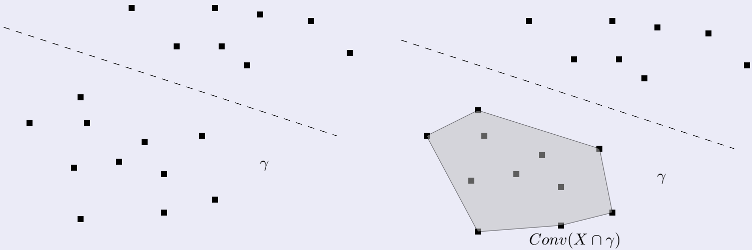
Instead of applying Helly's theorem to γ, we apply it to convex set $Conv(X \cap \gamma)$ which is compact.



Convex sets and centerpoints

Continuation of proof.

Letting γ run through all open half-spaces such that $|\gamma \cap X| > \frac{d}{d+1}|X|$ we obtain sets $Conv(X \cap \gamma)$ which do contain more than $\frac{d}{d+1}$ points of X. Because X is finite, there are only finitely many of such sets.



Convex sets and centerpoint

Continuation of proof.

Now we claim, that intersection of every $d+1$ sets from these convex sets is non-empty, and here's why:
Take $d+1$ sets: $C_1, \ldots, C_{d+1}$. We know, that C_i contains $(\frac{d}{d+1} + \varepsilon_i)$ points of X where $\varepsilon_i > 0$.

$$1 \geq |C_1 \cup C_2| = |C_1| + |C_2| - |\mathcal{I}_2| > \frac{2d}{d+1} - |\mathcal{I}_2| \Rightarrow |\mathcal{I}_2| > \frac{d-1}{d+1}$$

$$1 \geq |\mathcal{I}_2 \cup C_3| = |\mathcal{I}_2| + |C_3| - |\mathcal{I}_3| > \frac{2d-1}{d+1} - |\mathcal{I}_3| \Rightarrow |\mathcal{I}_3| > \frac{d-2}{d+1}$$

$$\ldots$$

$$1 \geq |\mathcal{I}_d \cup C_{d+1}| = |\mathcal{I}_d| + |C_3| - |\mathcal{I}_{d+1}| > \frac{1+d}{d+1} - |\mathcal{I}_{d+1}| \Rightarrow |\mathcal{I}_{d+1}| > 0$$

□

Introduction cd

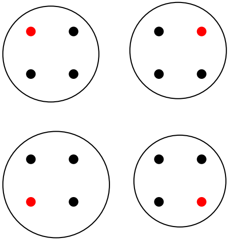
We've proved that the centerpoint always exists. Besides, it's known that the constant $\frac{2}{3}$ is the best possible.
Can we improve this constant by using, say, two points, or some other small number of points?
What happens when we replace convex sets by, say, axis-parallel rectangles?
Here we answer such questions.

Epsilon-nets

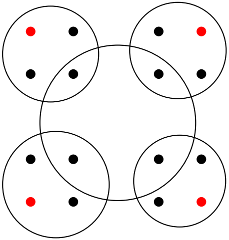
Definition 1

Let P be an n-point set in $\mathbb{R}^2$. Consider a family $\mathcal{S}$ of sets in $\mathbb{R}^2$. A set $Q \subset \mathbb{R}^2$ is called a weak $\varepsilon -$ *net* for P with respect to $\mathcal{S}$, if for any $S \in \mathcal{S}$ with $|S \cap P| > \varepsilon n$, we have $S \cap Q = \emptyset$. Further, if $Q \subseteq P$, then Q is called a (strong) ε -net for P with respect to $\mathcal{S}$

Example 1



Points - set P
Circles with interior - set F
Red points - strong $\frac{1}{4}$-net on the left example, but not on the right



What would it be about

Questions

- What's the minimal size of strong/weak epsilon-net for any $(P, \mathcal{S})$
- On which properties does bound depend?

This presentation

Let $0 \leq \varepsilon_i^{\mathcal{S}} \leq 1$ denote the smallest real number such that for any finite point set $P \subset \mathbb{R}^2$ there exist i-point set, which is $\varepsilon_i^{\mathcal{S}}$-net for P with respect to $\mathcal{S}$ ($\mathcal{S}$ is fixed).

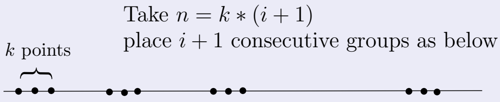
We try to obtain best bounds for $\varepsilon_i^{\mathcal{S}}$ for small values of i and family $\mathcal{S}$ being set of all axis-parallel rectangles, disks, half-planes and convex sets. We consider only set of points P in general position

General bounds

Theorem (Lemma 2.1)

If there exists a line L in the plane with the property that for every line segment on L there is a set $s \in \mathcal{S}$ such that $s \cap L$ is that segment, then $\varepsilon_i^{\mathcal{S}} \geq \frac{1}{i+1}$

Proof.



For such placement, if we assume $\varepsilon_i^{\mathcal{S}} < \frac{1}{i+1}$, each group has to contain one point from the net, hence $(i + 1)$ points are needed, a contradiction □

Theorem (Lemma 2.2)

If $\mathcal{S} \subset \mathcal{S}'$ then $\varepsilon_i^{\mathcal{S}} \leq \varepsilon_i^{\mathcal{S}'}$

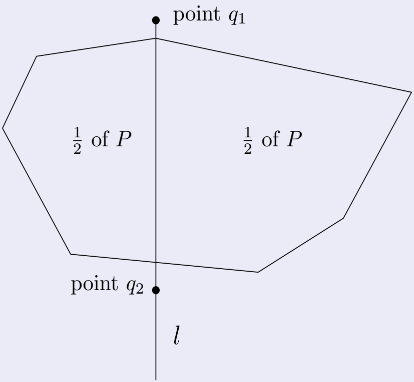
Half-planes

Let $\mathcal{H}$ denote the family of all half-planes.

Theorem (Lemma 2.3)

$\varepsilon_1^{\mathcal{H}} = \frac{2}{3}$, $\varepsilon_2^{\mathcal{H}} = \frac{1}{2}$, $\varepsilon_i^{\mathcal{H}} = 0$ *for* $i \geq 3$

Proof.



Let l be bisecting line of P. Any half-plane containing at least $1/2$ points of P must contain one of the points q_1, q_2. This proves $\varepsilon_2^{\mathcal{H}} \leq \frac{1}{2}$. On the other hand, for any n-point set and any points q_1, q_2, one of the two half-planes delimited by line going through $q_1 q_2$ contains at least $\frac{n-2}{2} = n(\frac{1}{2} - \frac{2}{n})$ points, so because $\frac{2}{n} \to 0$ we have $\varepsilon_2^{\mathcal{H}} \geq \frac{1}{2}$

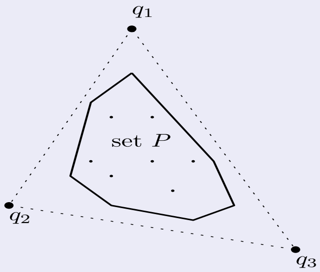
Half-planes

Let $\mathcal{H}$ denote the family of all half-planes.

Theorem (Lemma 2.3)

$\varepsilon_1^{\mathcal{H}} = \frac{2}{3}$, $\varepsilon_2^{\mathcal{H}} = \frac{1}{2}$, $\varepsilon_i^{\mathcal{H}} = 0$ *for* $i \geq 3$

Proof.



Given any point set P , pick $Q = \{q_1, q_2, q_3\}$ so that the triangle formed by those three points contains P . Thus any half-plane containing any point from P must contain at least one point of Q . This proves $\varepsilon_i^{\mathcal{H}} = 0$ for $i \geq 3$ □

Convex sets

Let $\mathcal{C}$ denote the family of all convex sets in the plane.

Theorem (Theorem 3.1)

$\varepsilon_2^{\mathcal{C}} \geq \frac{5}{9}$, $\varepsilon_3^{\mathcal{R}} = \frac{5}{12}$

Proof for 2-point net.

In order to prove lower bound, we need to...

Convex sets

Let $\mathcal{C}$ denote the family of all convex sets in the plane.

Theorem (Theorem 3.1)

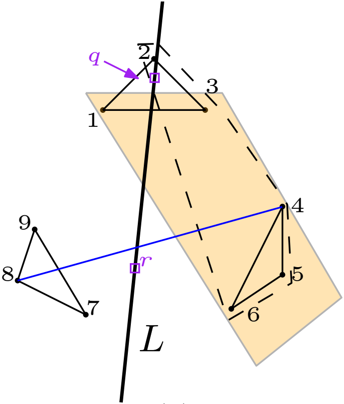
$\varepsilon_2^{\mathcal{C}} \geq \frac{5}{9}$, $\varepsilon_3^{\mathcal{R}} = \frac{5}{12}$

Proof for 2-point net.

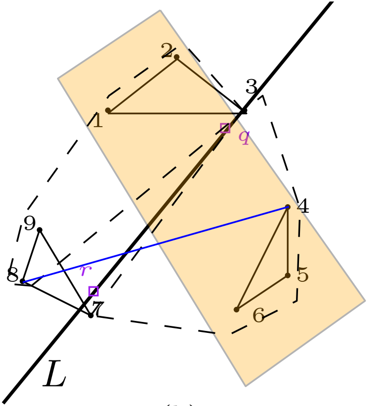
In order to prove lower bound, we need to construct set P of n points (for any n) s.t. for any pair (p, q) of points there exists convex set K which contains at least $5n/9$ of the points of P and avoids p, q.

Convex sets

The set P is made up of three groups, each consists of three subsets, arranged into a triangular shape. Each small subset, call them 1,2, ...,9 lies in some disk of some small diameter δ and contains $n/9$ points.



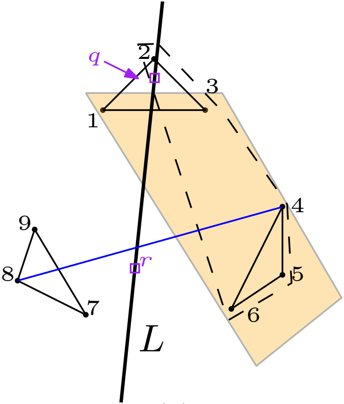
(a)



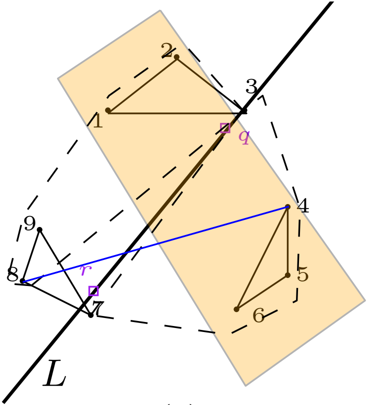
(b)

Convex sets

For any choice of q and r let L be the line through q and r.
Observe that L can intersect the convex hull of at most two of the subsets 1,...,9. We may assume, that L intersects at least one convex hull of some subset (otherwise we would already have $6n/9$ points lyinf on some side of L).



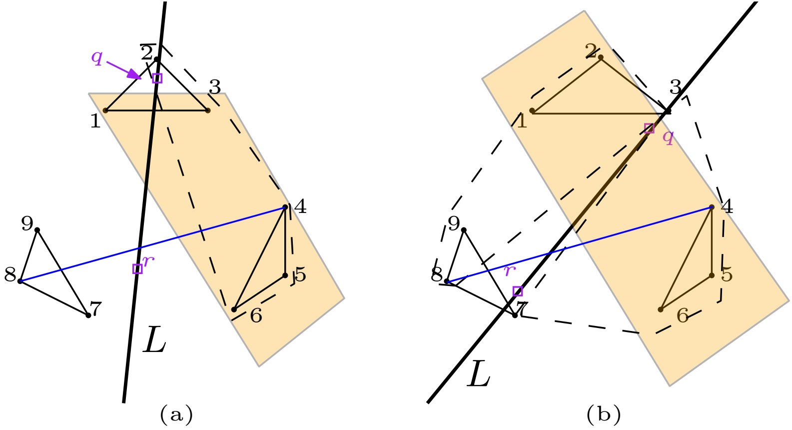
(a)



(b)

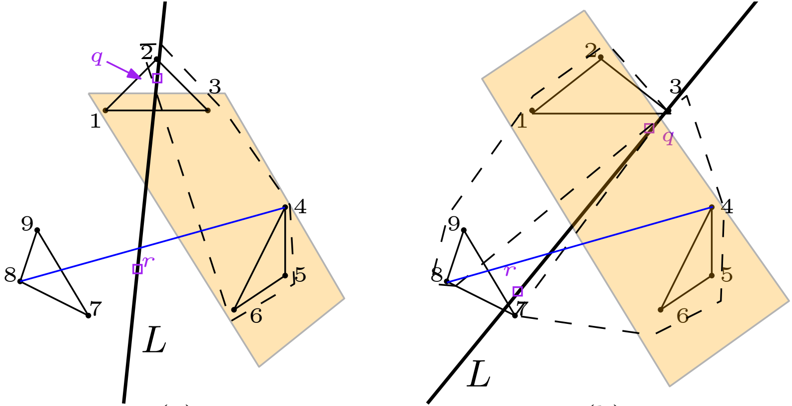
Convex sets

Moreover, we may assume, that L has at least 3 subsets fully lying on each sides. Otherwise, because L can "cross" at most 2 out of 9, we would have at least $9 - 2 - 2 = 5$ out of 9 subsets fully contained in one of the half-planes defined by L.



Convex sets

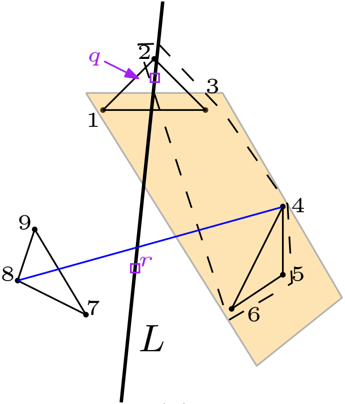
We write $CH(i, j, \dots)$ for convex hull of subsets $i \cup j \cup \dots$
WLOG assume L intersects $CH(1, 2, 3)$. Consider 2 cases:
a) L intersects $CH(2)$
b) L intersects $CH(3)$ (symmetrically, L intersects $CH(4)$)



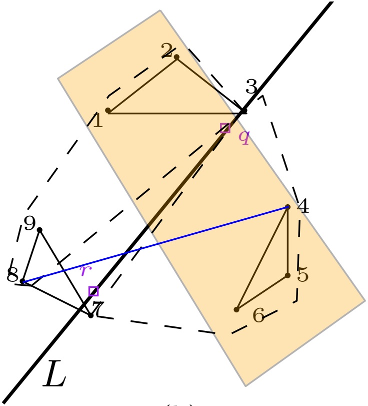
(a) (b)

Theorem 3.1 - case a, L intersects $CH(2)$

Exploiting symmetries, we can assume wlog that L is no closer to 6 than to 7. Then, in order to stab $CH(4,5,6,7,8)$, one of the points of Q has to lie on or below the upper tangent of $CH(4)$ and $CH(8)$.



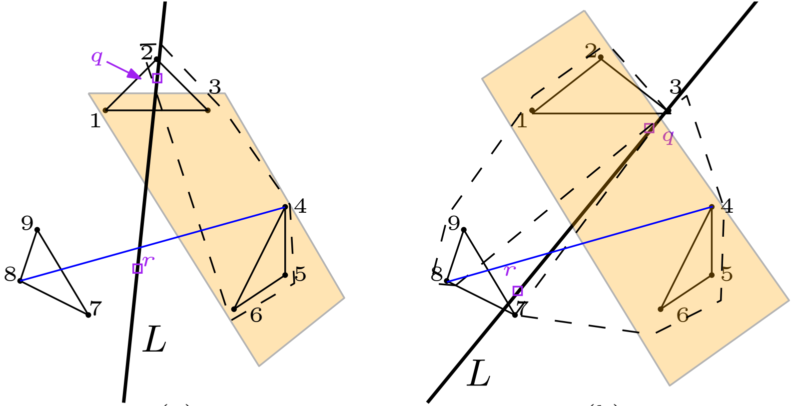
(a)



(b)

Theorem 3.1 - case a, L intersects $CH(2)$

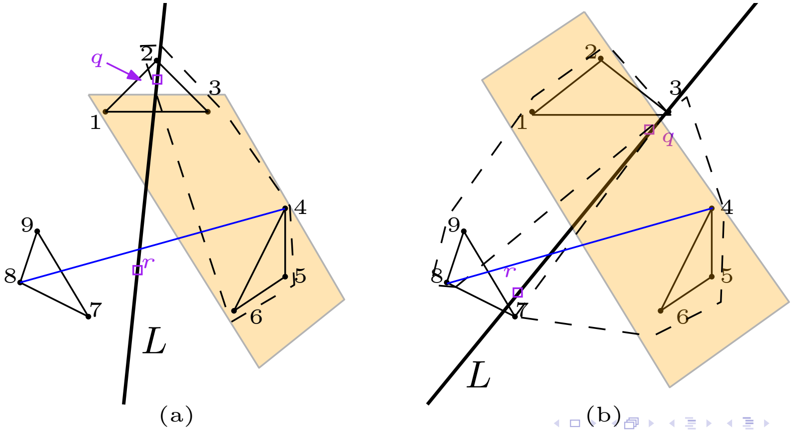
Since we must also have $q \in CH(2,3,4,5,6)$, q must lie arbitrarily close to 2 because disk containing all points from 2 can become sufficiently small. Therefore, for proper choice of δ our K would be $CH(1,3,4,5,6)$ and it avoids q and r



(a) (b)

Theorem 3.1 - case b, L intersects $CH(3)$

Observe, that in order to stab $CH(4,5,6,7,8)$ one of the points of Q must lie on or above the upper tangent of $CH(8)$ and $CH(4)$. If L is not closer to 8 than to 7, then we need $q \in L \cap CH(1,2,3,8,9)$. Otherwise, we need $q \in L \cap CH(3,4,5,6,7)$. In both cases q must lie atbitrarily close to $CH(3)$ if δ is chosen sufficiently small. Then $K = CH(1,2,4,5,6)$.



Theorem 3.1 - reminder

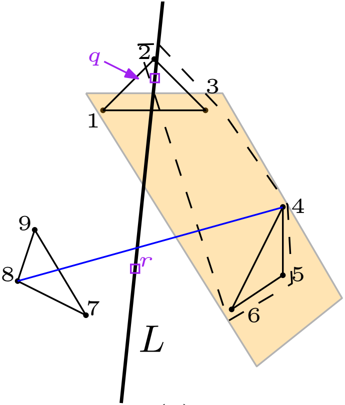
Let $\mathcal{C}$ denote the family of all convex sets in the plane.

Theorem (Theorem 3.1)

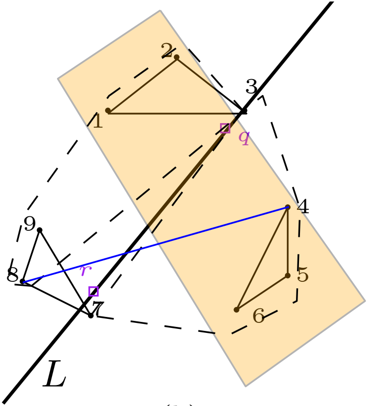
$\varepsilon_2^{\mathcal{C}} \geq \frac{5}{9}$, $\varepsilon_3^{\mathcal{C}} = \frac{5}{12}$

Theorem 3.1

To summarize, using our construction of point set P, for any 2 given points we can find a convex set K which avoids these points and contains $5n/9$ points from P. Thus, $\varepsilon_2^{\mathcal{C}} \geq \frac{5}{9}$.



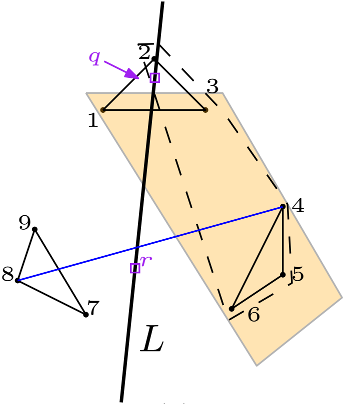
(a)



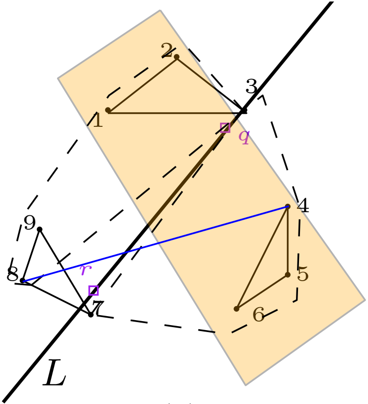
(b)

Theorem 3.1 for 3-point net

Let's examine our consruction for 2-point net at a higher level. We needed a "tangent condition" for point r and "closeness condition" for point q. We now place 4 triangular shaped groups (instead of the three) in a circular manner, each group consisting of three subsets of $n/12$ points. This gives $\binom{4}{3} = 4$ instances of type before.



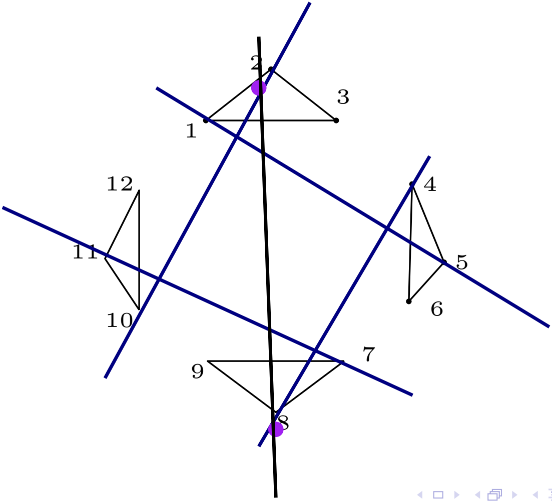
(a)



(b)

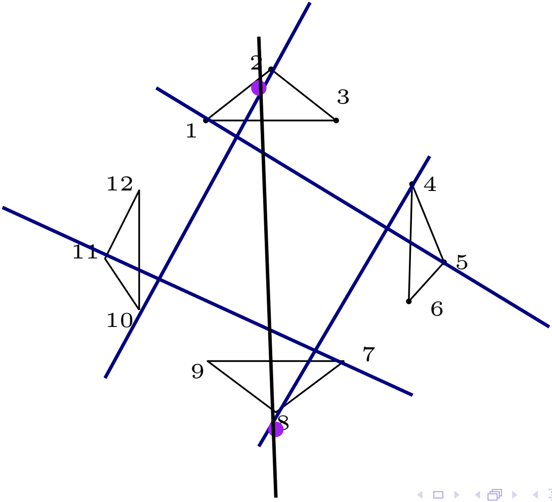
Theorem 3.1 for 3-point net

Because we have 4 instances of type before, we need to satisfy 4 tangent conditions and 4 closeness conditions. Two points suffice to satisfy all the tangent conditions. Still 4 "closeness conditions" left



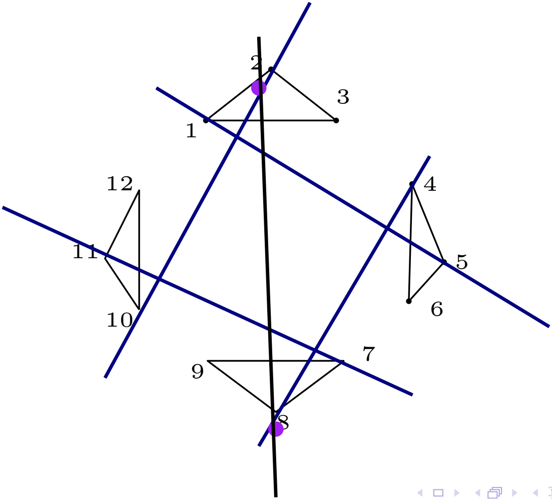
Theorem 3.1 for 3-point net

Because we have 4 instances of type before, we need to satisfy 4 tangent conditions and 4 closeness conditions. Two points suffice to satisfy all the tangent conditions plus two closeness conditions.



Theorem 3.1 for 3-point net

However, the third point cannot satisfy two other closeness conditions simultaneously. Hence, we can also construct a convex set with 5 parts which would have $5n/12$ points and avoid any 3-point net.



Convex sets - upper bounds

Theorem (Ham-Sandwich theorem)

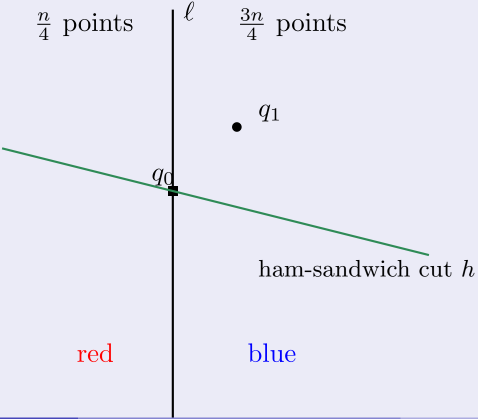
Every d finite sets in $\mathbb{R}^d$ can be simultaneously bisected by a hyperplane. A hyperplane bisects set A if each open half-space defined by that hyperplane contains at most $\lceil \frac{|A|}{2} \rceil$ points of A

Convex sets - upper bounds

Theorem (Theorem 3.2)

$$\varepsilon_2^{\mathcal{C}} \geq \frac{5}{8}, \varepsilon_3^{\mathcal{C}} = \frac{7}{12}, \varepsilon_4^{\mathcal{C}} = \frac{4}{7}, \varepsilon_5^{\mathcal{C}} = \frac{1}{2},$$

Proof for 2-point net.

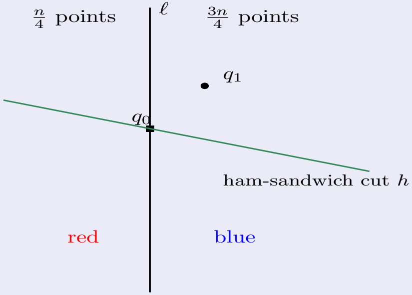


Convex sets - upper bounds

Theorem (Theorem 3.2)

$$\varepsilon_2^{\mathcal{C}} \leq \frac{5}{8}, \varepsilon_3^{\mathcal{C}} \leq \frac{7}{12}, \varepsilon_4^{\mathcal{C}} \leq \frac{4}{7}, \varepsilon_5^{\mathcal{C}} \leq \frac{1}{2},$$

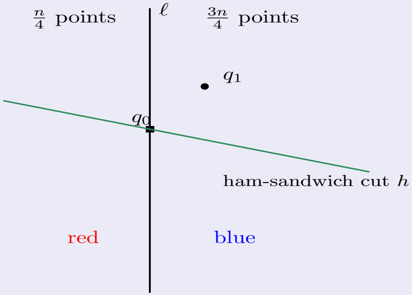
Proof for 2-point net.



Let q_1 be the centerpoint for blue points. q_0 is defined as $\ell \cap h$. Let K be any convex set with $q_0, q_1 \notin K$. As $q_0 \notin K$, the set K avoids at least one of the four quadrants defined by ℓ and h (by convexity).

Convex sets - upper bounds

Proof for $\varepsilon_2^{\mathcal{C}} \leq \frac{5}{8}$.



If this quadrant is blue then K avoids at least $3n/8$ (blue) points, if it's red then K avoids at least $n/8$ (red) points. In addition, because $q_1 \notin K$ and q_1 is "blue center-point", K avoids at least $\frac{1}{3} \cdot \frac{3n}{4} = \frac{n}{4}$ blue points. Altogether K avoids at least $3n/8$ points, so in either case K can't contain more than $5n/8$ points. Other proofs are similar, line ℓ is chosen differently. □

Convex sets - upper bounds cd

One could recursively apply constructions as above, which leads to bound $\varepsilon_i^{\mathcal{C}} \leq \frac{2}{3}\left(\frac{3}{4}\right)^k$ for $i = \frac{1}{3}(4^{k+1} - 1), k \geq 0$
A rough calculation shows that a weak ε-net of size $\mathcal{O}(\frac{1}{\varepsilon^5})$ with respect to $\mathcal{C}$ is obtained. Unfortunately it falls short of the best known bound $\mathcal{O}(\frac{1}{\varepsilon^2})$.
Still, these constructions are better for small nets.

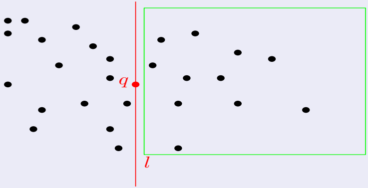
Axis-parallel rectangles

Let $\mathcal{R}$ denote the family of all axis-parallel rectangles.

Theorem

$\varepsilon_1^{\mathcal{R}} \geq \frac{1}{2}$, $\varepsilon_2^{\mathcal{R}} = \frac{2}{5}$, $\varepsilon_3^{\mathcal{R}} \geq \frac{2}{6}$

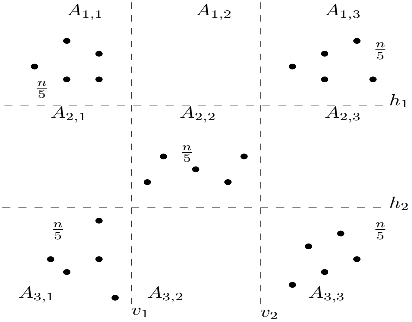
Proof for 1-point net.



Given any point set P and any point q, we can also construct a rectangle which contains at least $\lfloor \frac{n-1}{2} \rfloor \geq n/2 - 2 = n(\frac{1}{2} - \frac{2}{n})$ points. Thus, $\varepsilon_1^{\mathcal{R}} \geq \frac{1}{2}$ because n can be chosen to be arbitrarily large.

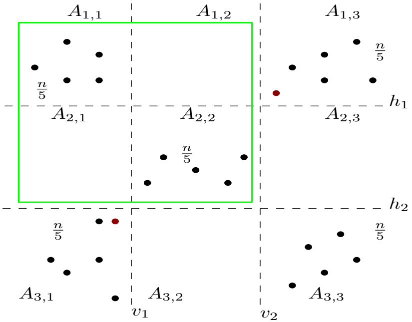
□

Axis-parallel rectangles-2.1



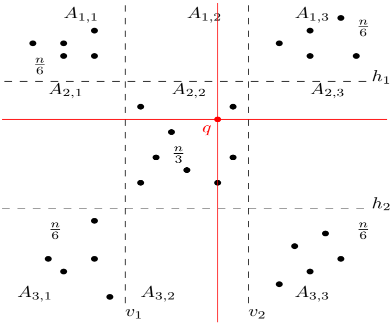
Suppose for contradiction that $\varepsilon_2^{\mathcal{R}} = \varepsilon < \frac{2}{5}$. If a pair of points Q $= \{q_1, q_2\}$ is a weak ε-net for P with respect to axis-parallel rectangles and $\varepsilon < 2/5$, then each of the four strips above h_1, below h_2, left of v_1 and right of v_2 must contain a point of Q . Since no triple of strips has a common intersection, each of the 2 points must be contained in exactly two strips. Then either $Q \subset A_{1,3} \cup A_{3,1}$ or $Q \subset A_{1,1} \cup A_{3,3}$. Assume wlog the former case.

Axis-parallel rectangles-2.2



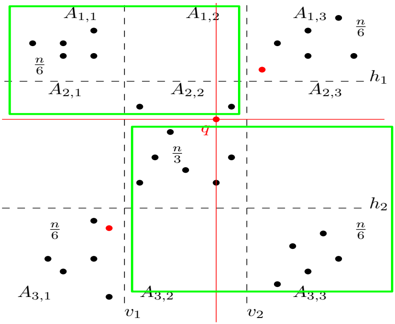
We've assumed $Q \subset A_{1,3} \cup A_{3,1}$. Let red points be points from Q. But then we can immediately construct green rectangle, containing $\frac{2}{5}n$ points and avoiding Q, a contradiction.

Axis-parallel rectangles-3.1



Next we prove $\varepsilon_3^{\mathcal{R}} \geq \frac{2}{6}$. Assume for contradiction $\varepsilon_3^{\mathcal{R}} = \varepsilon < \frac{2}{6}$. First, observe that one point from Q should be inside $A_{2,2}$. Let this point be q. Next, by argument from previous proof, we claim that two other points of Q must be either in $A_{1,1} \cap A_{3,3}$ or in $A_{1,3} \cap A_{3,1}$. Assume latter case wlog.

Axis-parallel rectangles-3.2

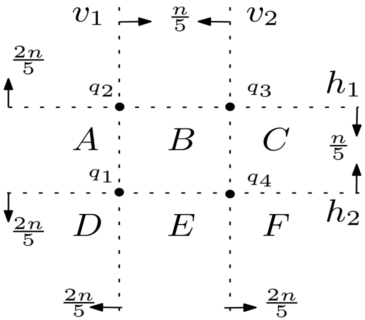


We've assumed $Q \subset A_{1,3} \cap A_{3,1} \cap A_{2,2}$. But now it's easy to see that one of the green rectangles must contain at least $\frac{n}{6} + \frac{n}{3 \cdot 2} - 1 = \frac{n}{3} - 1$ points, and both are avoiding Q. Since for n large enough, $\frac{1}{3} - \frac{1}{n} > \varepsilon$ we have a contradiction.

Axis-parallel rectangles

Theorem (Theorem 4.3)

$\varepsilon_3^{\mathcal{R}} \leq \frac{2}{5}$.

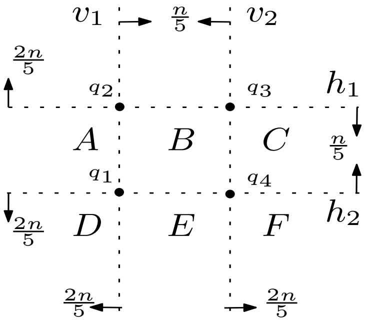


Let v_1 be a vertical line with exactly $2/5 \cdot n$ points of P to and let v_2 be a vertical line with exactly $2/5 \cdot n$ points of P to its right. Similarly consider a line h_1 (resp., h_2) with exactly $2/5 \cdot n$ points of P above it (resp., below it).Let $\{q_1, \ldots, q_4\}$ be points of intersection of these lines.

Axis-parallel rectangles

Theorem (Theorem 4.3)

$\varepsilon_3^{\mathcal{R}} \leq \frac{2}{5}$.

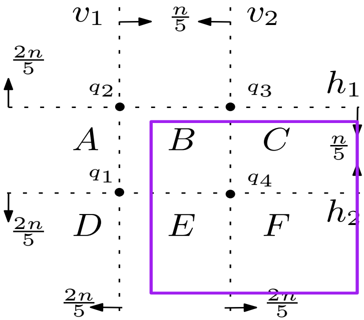


Observe that $Q = \{q_1, \ldots, q_4\}$ is $\frac{2}{5}$ -net for P. Let $Q_1 = \{q_1, q_3\}$, $Q_2 = \{q_2, q_4\}$. We'll show that at least one of Q_1, Q_2 is a 2-point $\frac{2}{5}$-net for P. Assume to the contrary that neither is.

Axis-parallel rectangles

Theorem (Theorem 4.3)

$\varepsilon_3^{\mathcal{R}} \leq \frac{2}{5}$.

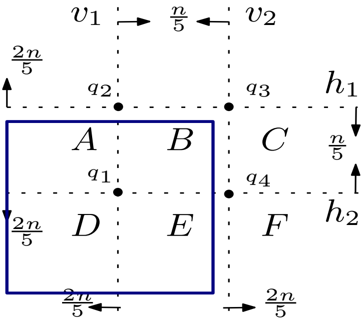


So Q_1 is not $\frac{2}{5}$-net for P. That means, there exist a rectangle containing more than $\frac{2}{5}$ points and avoiding q_1, q_3. Observe that such rectangle should contain either q_2 or q_4. Assume wlog it contains q_4. Symmetrically, there must exist a rectangle proving that Q_2 is not a weak $\frac{2}{5}$-net, and suppose it contains q_1

Axis-parallel rectangles

Theorem (Theorem 4.3)

$\varepsilon_3^{\mathcal{R}} \leq \frac{2}{5}$.

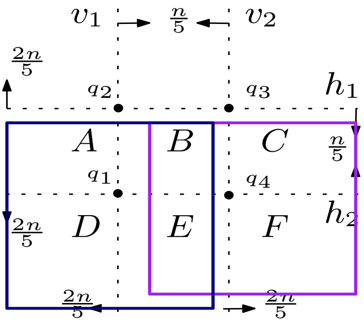


Symmetrically, there must exist a rectangle proving that Q_2 is not a weak $\frac{2}{5}$-net, and suppose it contains q_1

Axis-parallel rectangles

Theorem (Theorem 4.3)

$\varepsilon_3^{\mathcal{R}} \leq \frac{2}{5}$.



Let $A, \ldots, F$ - amount of points inside corresponding rectangles, induced by lines h_1, h_2, v_1, v_2 (**not** inside colored rectangles). Now we have:

$$A + B + D + E > \frac{2n}{5}$$

$$B + C + E + F > \frac{2n}{5}$$

$$A + B + C = \frac{n}{5}, \; D + E + F = \frac{2n}{5}$$

$$\Rightarrow B + E > \frac{n}{5},$$

a contradiction, which ends the proof.

Axis-parallel rectangles - general lemma

Theorem (Lemma 4.2)

For all positive integers k, i, j *and* $\ell \leq k+1$,
$\varepsilon^{\mathcal{R}}_{k^2+2\ell i+2(k+1-\ell)j} \leq \frac{\varepsilon^{\mathcal{R}}_i \varepsilon^{\mathcal{R}}_j}{\ell \varepsilon^{\mathcal{R}}_j + (k+1-\ell)\varepsilon^{\mathcal{R}}_i}$.

Using this lemma, we can obtain the following bounds:

$$\varepsilon^{\mathcal{R}}_1 \leq \frac{1}{2}, \varepsilon^{\mathcal{R}}_3 \leq \frac{1}{3}, \varepsilon^{\mathcal{R}}_5 \leq \frac{1}{4}, \varepsilon^{\mathcal{R}}_7 \leq \frac{2}{9}, \varepsilon^{\mathcal{R}}_8 \leq \frac{1}{5}, \varepsilon^{\mathcal{R}}_{10} \leq \frac{1}{6}, \varepsilon^{\mathcal{R}}_{16} \leq \frac{2}{15}$$

Remark 1 - disks on the plane

Theorem (Theorem 5.1)

It is interesting to note that some bounds on the size of weak ε-nets follow rather directly from classical results. We illustrate this fact for the collection $\mathcal{D}$ of all disks in the plane. $\varepsilon_4^{\mathcal{D}} \leq \frac{1}{2}$.

Let P be a set of n points in the plane. We need to show that there exists a set Q of four points such that every disk d for which $|d \cap P| > \frac{n}{2}$ must intersect Q. Consider the collection $D \subset \mathcal{D}$ of all disks d that contain more than $n/2$ points of P. Obviously every pair of disks of D must have a non-empty intersection. By the result of [6], there exists a set Q of four points that stab all disks in D . This completes the proof.

Remark 2 - results

	Convex sets		Half-planes		Disks		Rectangles	
	LB	UB	LB	UB	LB	UB	LB	UB
ε_0	1		1		1		1	
ε_1	2/3		2/3		2/3		1/2	
ε_2	5/9	5/8	1/2		1/2	5/8	2/5	
ε_3	5/12	7/12	0		1/4	7/12	1/3	
ε_4	1/5	4/7	0		1/5	1/2	1/5	5/16
ε_5	1/6	1/2	0				1/6	1/4

It's been shown that $\varepsilon_i^{\mathcal{R}} \leq \frac{2}{i+3}$ for all $1 \leq i \leq 5$. It's open whether it holds for all i.

One hypothesis is that it's true, for nets chosen from grid similar as appeared in previous proofs.