EXERCISE

① Voting

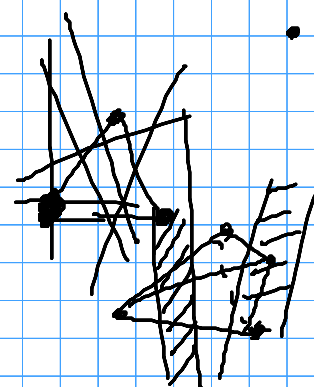
$1 - d = 0.8 \rightarrow d = 0.2$

voter: 62 millionen

$\frac{1}{\varepsilon} \cdot \ln 5$ $\qquad \frac{1}{d} = 5$

		ε
CDU	39,5 %	0.395
SPD	26,0 %	0.26
GRÜNE	10,5 %	0.105
FDP	5,0 %	0.05
LINKE	8,0 %	0.08

② U - points in $\mathbb{R}^n$



$\mathcal{S}$ - hyperplanes

t.s: VC-dim $\leq n+1$

lower bound: $\exists$ set of $n+1$ points that can be shattered

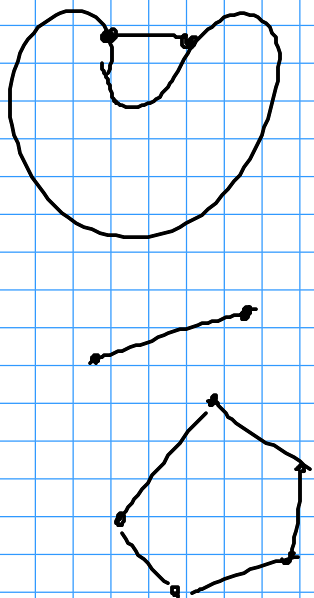
regular simplex

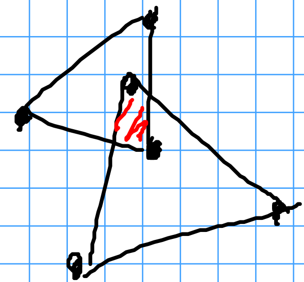
↳ corners are unit vector
$(0, \ldots, 0)$
$(1, 0, \ldots, 0)$
$(0, 1, 0, \ldots, 0)$
$\vdots$

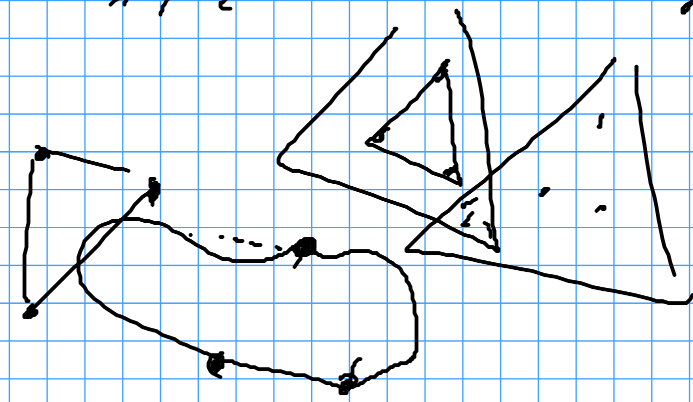
$\Rightarrow$ every subset forms a facet of the simplex $\rightarrow$ cut away by a hyperplane

upper bound: $n+2$

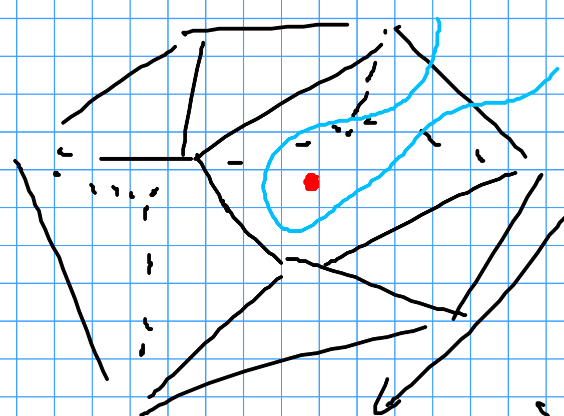
Radon's theorem to show that S_1 can not be shattered, where $S_1 \subseteq S$ and S_1, S_2 have intersecting convex hulls







$\Rightarrow$ as soon as $CH(S_1)$ and $CH(S_2)$ intersect $\Rightarrow$ at least one point from S_2 is inside $CH(S_1)$



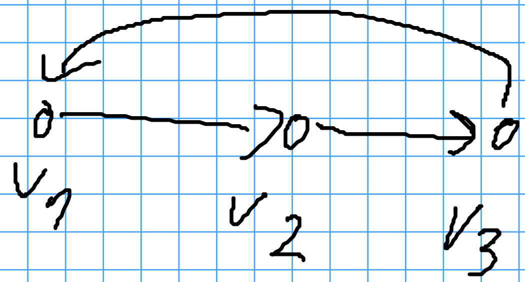
points of S_1 which form the convex hull can not be separated from points of S_2 inside $CH(S_1)$

$\Rightarrow$ S_1 can not be separated from S_2

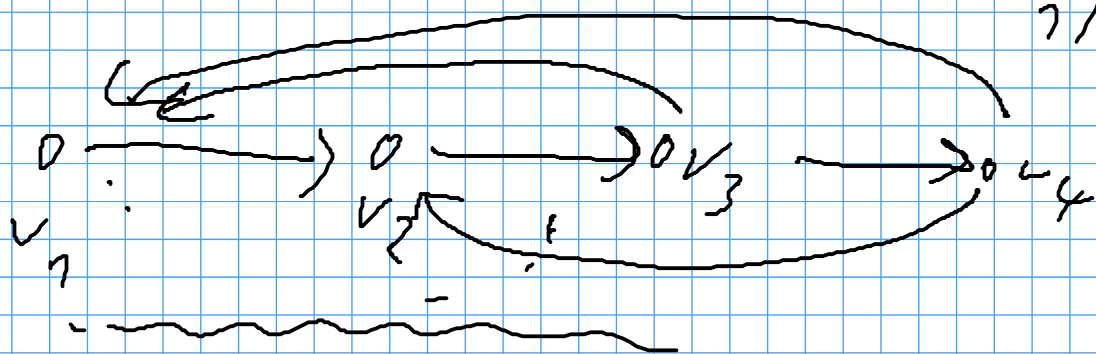
$\Rightarrow$ S can not be shattered

because hyperplanes are convex

③ VC-dim. of unique directed shortest path systems



v_1, v_2, v_3, v_4



~~v_1 v_2 v_4~~

~~v_1 v_4 v_2~~

~~v_2 v_1 v_4~~

~~v_2 v_4 v_1~~

~~v_4 v_1 v_2~~

~~v_4 v_2 v_1~~

v_1 v_2

v_2 v_3

v_3 v_4

v_1 v_2 v_3

v_2 v_3 v_4

~~v_1 v_4~~

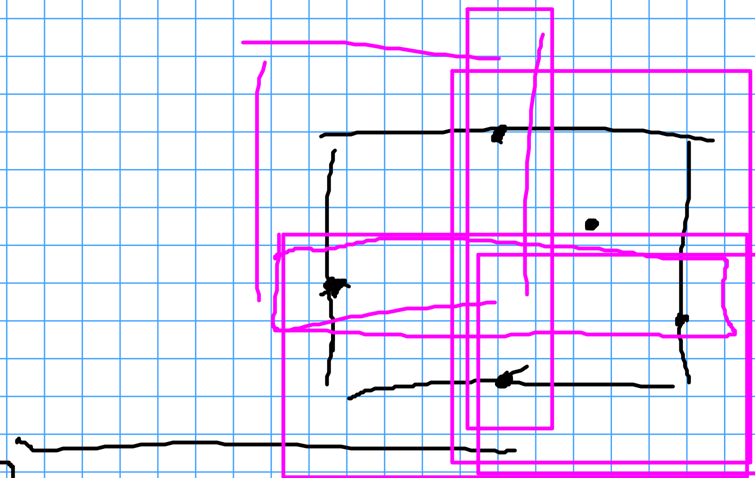
~~v_1 v_3~~

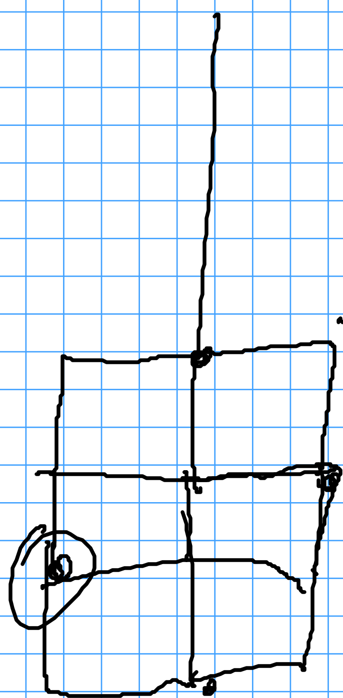
~~v_2 v_4~~

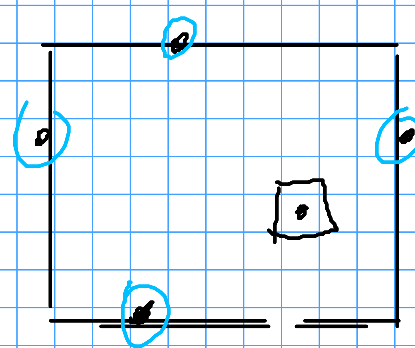
~~v_1 v_3 v_4~~

v_1 v_2 v_4

④ VC-dim axis parallel rectangles = 4







Smallest enclosing axis-parallel rectangle

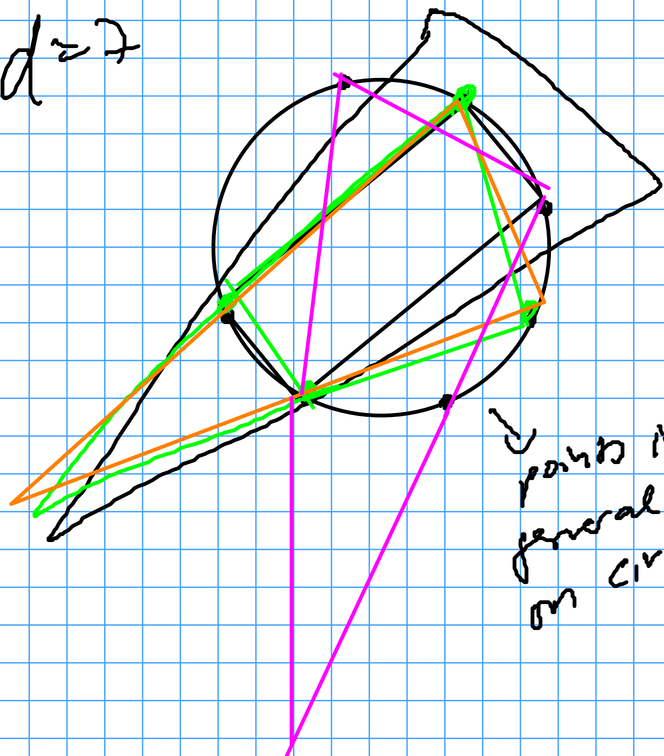
⇒ formed by four points

⇒ 5^{th} point inside this rectangle

⇒ four points around can not be separated from the one inside

VC-dim of triangles

$d = 7$

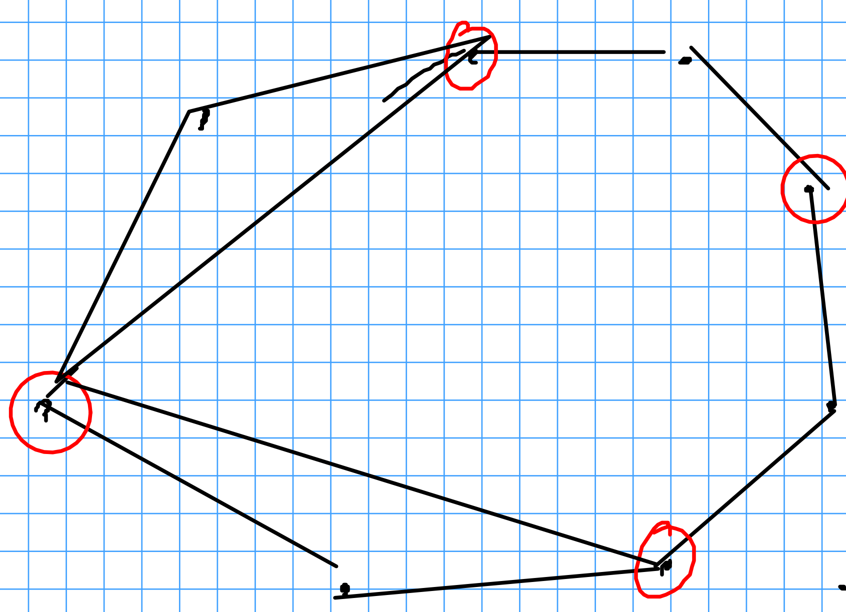


points in general position on circle

Shatter 1 point ✓
Shatter 2 thin triangle
Shatter 3 triangle
Shatter ≥ 4 → cut away 3 or less points is always possible as triangle allows for 3 straight cuts

Why not shatter 8 points?

if not all points on CH then CH-points can not be shattered



→ consider every 2^{nd} point on CH

→ have to cut away four points with three straight cuts

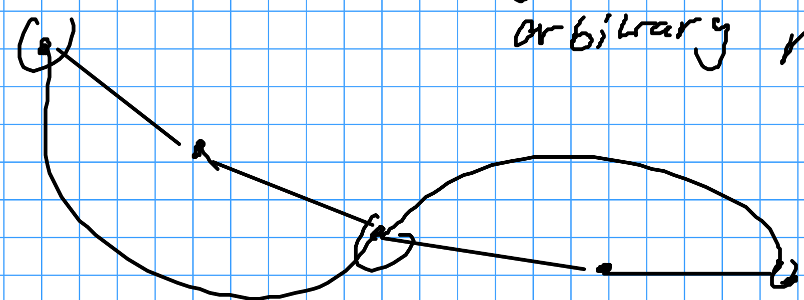
→ no two points can be cut away with one cut

↳ triangle can not shatter this set

$\Longrightarrow d \leq 7$

unbounded VC-dim?

→ graph with arbitrary paths



→ all polygons with arbitrary number of corners

→ in general $|\mathcal{S}| = 2^n$

fix VC-dim d $|\mathcal{S}| = n^d$

Def. (weighted ε-net)

Let $(U, \mathcal{S})$ be a set system, $w: \mathcal{S} \to \mathbb{R}^+$ a weight function, and $\varepsilon \in [0,1]$. A set $R \subseteq U$ is called an ε-net for $(U, \mathcal{S})$, if $\forall S \in \mathcal{S}$ with $w(S) \geq \varepsilon \cdot w(U)$: $R \cap S \neq \emptyset$.

normal ε-net : $\omega(S) = |S|$

2.5.5. Approximation Alg. Based on Small VC-dim. and ε-nets

Basic ε-Net Theorem

- Let $(U, \mathcal{S})$ be a set system with VC-dim. d, and $\varepsilon, \delta \in [0,1]$ parameters. A random sample of size

$$\frac{d}{\varepsilon} \log \frac{d}{\varepsilon} + \frac{1}{\varepsilon} \log \frac{1}{\delta}$$

is an ε-net with probability $\geq \delta$.

- Let $(U, \mathcal{S})$ be a set system with VC-dim $d \geq 2$ and $\varepsilon \in]0,1[$ a parameter, then there exists an ε-net of size $O\left(\frac{d}{\varepsilon} \log \frac{1}{\varepsilon}\right)$ for $(U, \mathcal{S})$.

↳ prove this by showing that a random sample of sufficient size is an ε-net with positive probability.

Small helping lemma

Lemma (Chernoff-Bound) Let $X = \sum_{i=1}^{n} X_i$ be a random variable, with X_i being independent of each other and $P(X_i = 1) = p$.

Then $P(X \geq \frac{np}{2}) \geq \frac{1}{2}$ if $np \geq 8$.

Proof. $P(|X - E(X)| > t\sigma) < \frac{1}{t^2}$ for $t > 0$

(Chebyshev's inequality)

For a binomial distribution we know that $E(X) = np$ and

$Var(X) = \sum_{i=1}^{n} Var(X_i) \leq np$. So we get

$$P(X < \frac{np}{2}) \leq P(|X - E(X)| > \frac{np}{2})$$

$$\leq \frac{4}{np} \underset{np \geq 8}{\leq} \frac{1}{2}$$

//

Exercise: • random graph creator

↳ random points in plane

↳ connect two points via an edge if their distance $\leq r$

↳ some alg. to get shortest paths in that graph

$\Rightarrow$ in the end, we want to have an alg. which sits all 'long' paths with a

Small number of vertices in this graph $\Leftrightarrow$ constructing an ε-net for the set system, with ε about $1/\log n$