# Journey to the Center of the Point Set

Sariel Har-Peled and Mitchell Jones

SoCG ’19, June 18, 2019

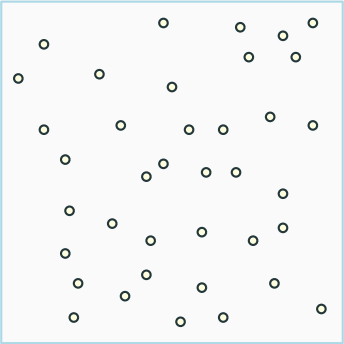
University of Illinois at Urbana-Champaign

Student Presentation

### Definition: Centerpoints

$P \subset \mathbb{R}^d$: set of $n$ points.

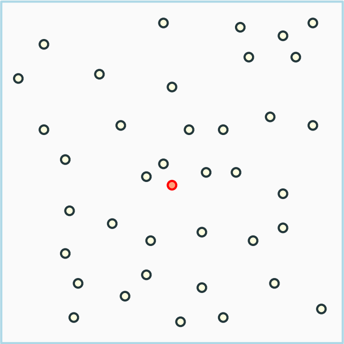
$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.



## Definition: Centerpoints

$P \subset \mathbb{R}^d$: set of $n$ points.

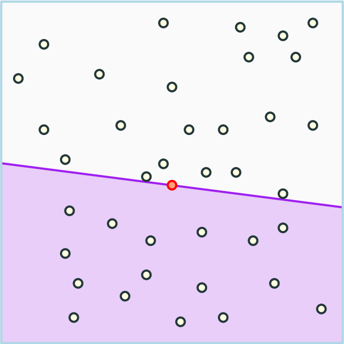
$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.



## Definition: Centerpoints

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.



### Definition: Centerpoints

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.

Applications:

**Definition: Centerpoints**

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.

Applications:

- One point summary of $P$

**Definition: Centerpoints**

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.

Applications:

- One point summary of $P$
- Divide and conquer

**Definition: Centerpoints**

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.

Applications:

- One point summary of $P$
- Divide and conquer
- Helly's Theorem $\implies$ existence

### Definition: Centerpoints

$P \subset \mathbb{R}^d$: set of $n$ points.

$c \in \mathbb{R}^d$ centerpoint for $P$ if for every closed halfspace $h^+$:
$c \in h^+ \implies |P \cap h^+| \geqslant n/(d+1)$.

Applications:

- One point summary of $P$
- Divide and conquer
- Helly's Theorem $\implies$ existence
- $\alpha$-centerpoints for $\alpha \in (0, 1/(d+1)]$

# Previous work: computing centerpoints

- Brute force $\Theta(n^d)$ time

# Previous work: computing centerpoints

- Brute force $\Theta(n^d)$ time
- $O(n^{d-1} + n \log n)$ expected time [Chan, 2004]

# Previous work: computing centerpoints

- Brute force $\Theta(n^d)$ time
- $O(n^{d-1} + n \log n)$ expected time [Chan, 2004]
- $\approx 3/4(d+2)^2$-centerpoint, randomized time $O\big((d^5 \log d)^{\log_2 d}\big)$ [Clarkson, Eppstein, Miller, Sturtivant, and Teng, 1996]

# Previous work: computing centerpoints

- Brute force $\Theta(n^d)$ time
- $O(n^{d-1} + n \log n)$ expected time [Chan, 2004]
- $\approx 3/4(d+2)^2$-centerpoint, randomized time $O\big((d^5 \log d)^{\log_2 d}\big)$ [Clarkson, Eppstein, Miller, Sturtivant, and Teng, 1996]
- Can be derandomized [Miller and Sheehy, 2010]

# Previous work: computing centerpoints

- Brute force $\Theta(n^d)$ time
- $O(n^{d-1} + n \log n)$ expected time [Chan, 2004]
- $\approx 3/4(d+2)^2$-centerpoint, randomized time $O\big((d^5 \log d)^{\log_2 d}\big)$ [Clarkson, Eppstein, Miller, Sturtivant, and Teng, 1996]
- Can be derandomized [Miller and Sheehy, 2010]
- **Open:** $\approx 1/(d+1)$-centerpoint in $O(\text{poly}(d))$ time?

# A polynomial algorithm

**Theorem** [Clarkson, Eppstein, Miller, Sturtivant, and Teng, 1996]

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $1/(4(d+2)^2)$-centerpoint in time $O(d^9 \log d)$.

# A polynomial algorithm

**Theorem** [Clarkson, Eppstein, Miller, Sturtivant, and Teng, 1996]

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $1/(4(d+2)^2)$-centerpoint in time $O(d^9 \log d)$.

**Our result**

With random sampling, $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

**Our result**

With random sampling, a $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

**Our result**

With random sampling, a $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

- Approximate centerpoint for $d + O(1)$ points in $\mathbb{R}^d$?

# The idea

**Our result**

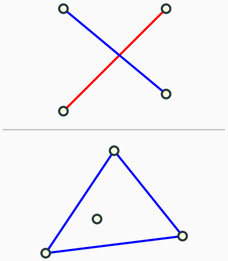
With random sampling, a $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

- Approximate centerpoint for $d + O(1)$ points in $\mathbb{R}^d$?
- Yes! Radon's Theorem.

### Radon's Theorem

$P \subset \mathbb{R}^d$: set of $d+2$ points.

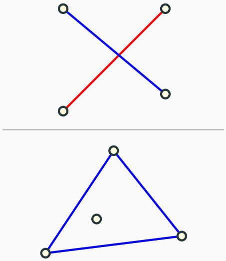
$\exists$ partition $P = Q \sqcup R$ s.t. $\text{conv}(Q) \cap \text{conv}(R) \neq \varnothing$.



**Radon's Theorem**

$P \subset \mathbb{R}^d$: set of $d+2$ points.

$\exists$ partition $P = Q \sqcup R$ s.t. $\text{conv}(Q) \cap \text{conv}(R) \neq \varnothing$.

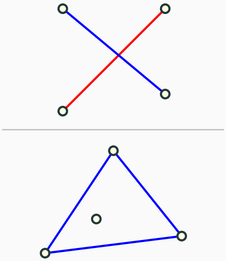


- Radon point: compute in $O(d^3)$ time.

**Radon's Theorem**

$P \subset \mathbb{R}^d$: set of $d + 2$ points.

$\exists$ partition $P = Q \sqcup R$ s.t. $\mathrm{conv}(Q) \cap \mathrm{conv}(R) \neq \varnothing$.



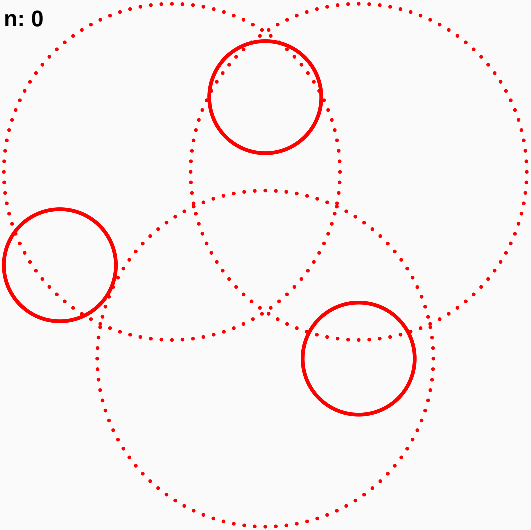
- Radon point: compute in $O(d^3)$ time.
- Radon point: $2/(d + 2)$-centerpoint for $P$.

# Our contribution

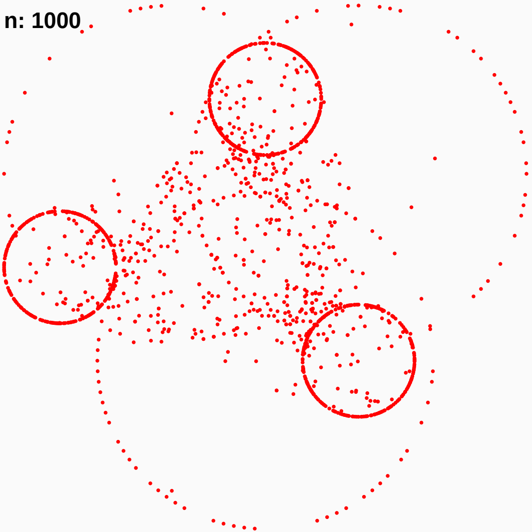
A simplified variant of [Clarkson, Eppstein, et al., 1996].

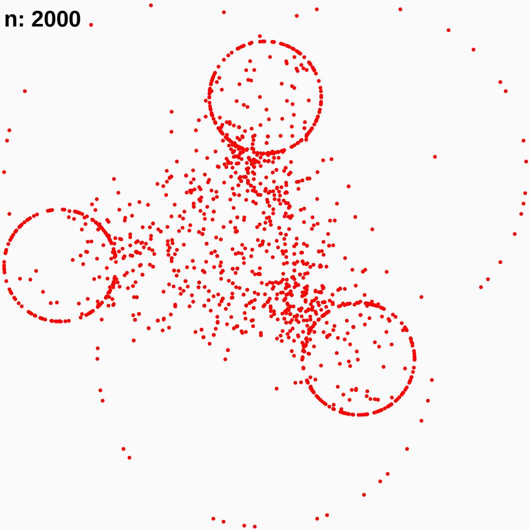
1. $Q \subseteq P$ sample of size $\approx O(d^3 \log d)$ [Li, Long, et al., 2001]
2. For $i = 1, \ldots, O(d|Q|)$:
   - 2.1 Sample $d + 2$ points of $Q$
   - 2.2 Compute their radon point $r$
   - 2.3 Add $r$ to $Q$
   - 2.4 Delete a random point from $Q$ (which isn't $r$)

# Visualization

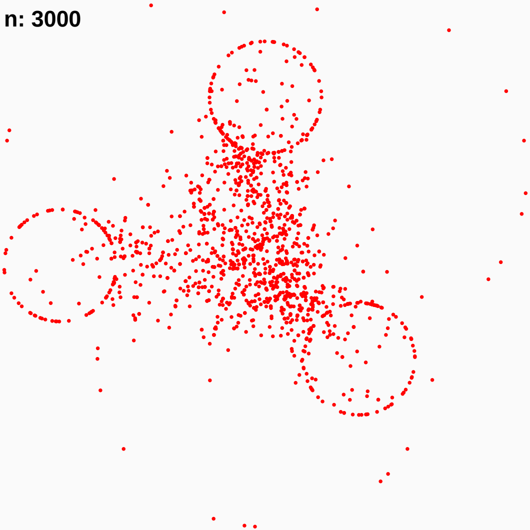


# Visualization

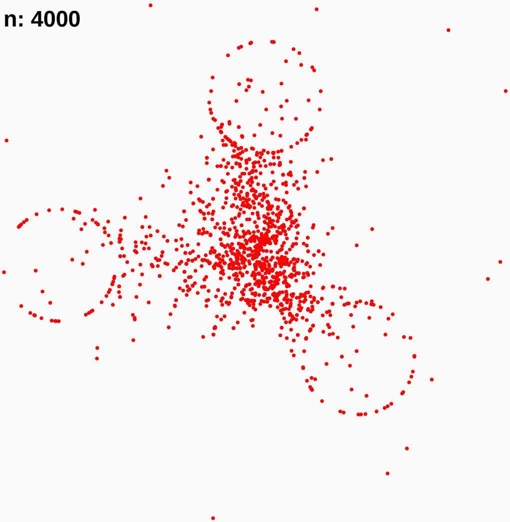



n: 2000

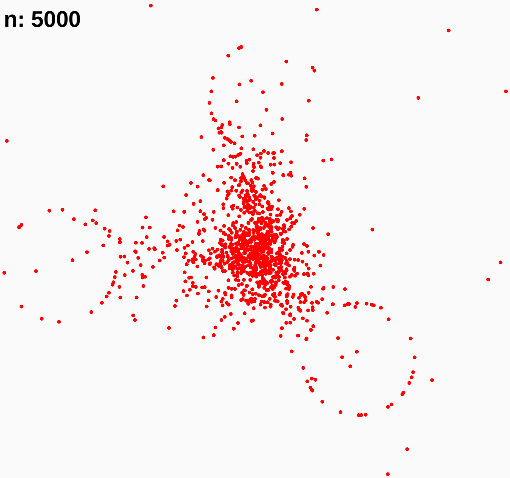
# Visualization



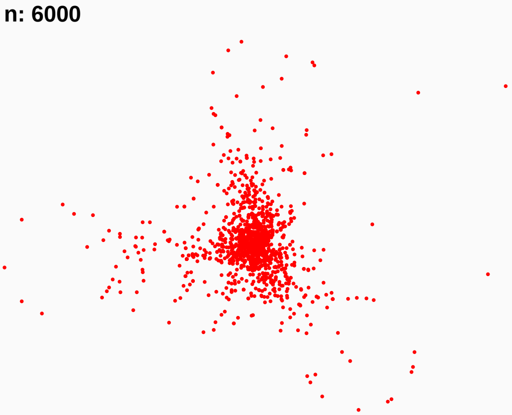
# Visualization



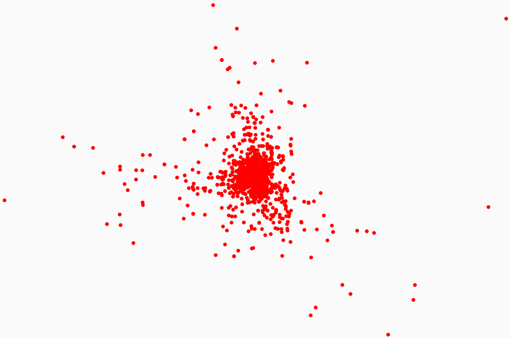
# Visualization



# Visualization

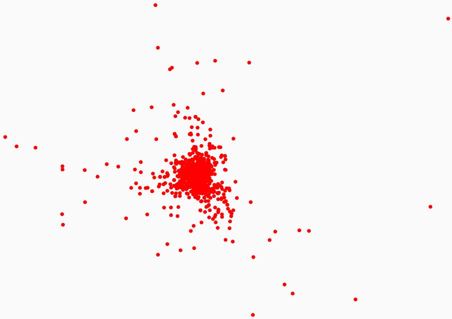


n: 7000

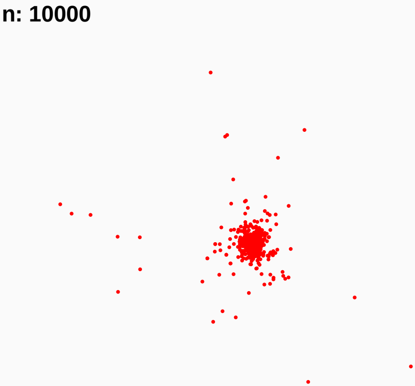


# Visualization

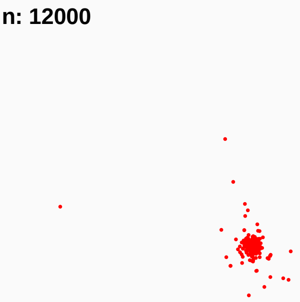
**n: 8000**

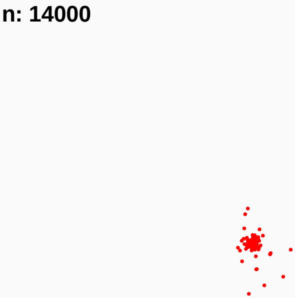


# Visualization



# Visualization




n: 14000

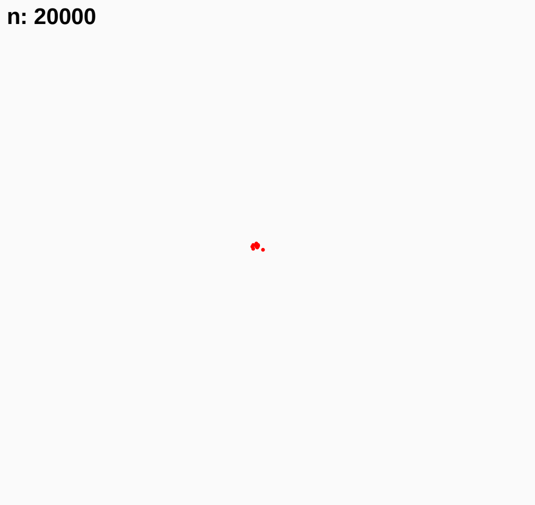
# Visualization

**n: 16000**

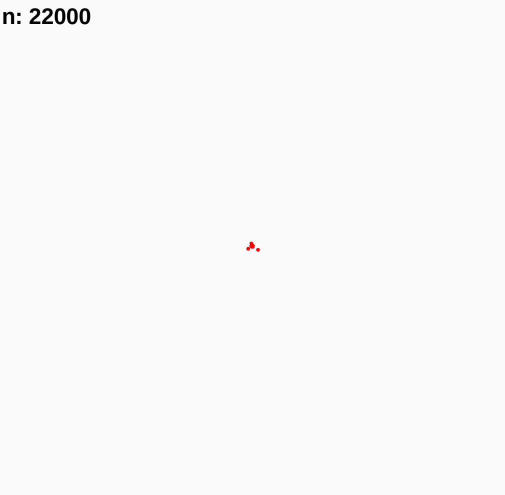
# Visualization

n: 18000

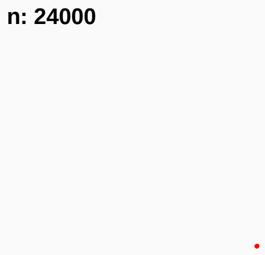
# Visualization

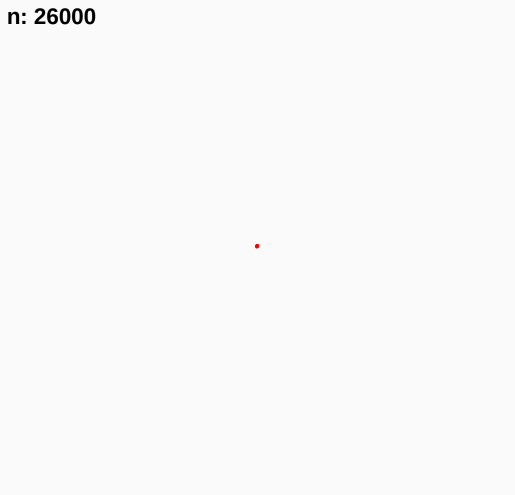


# Visualization



# Visualization



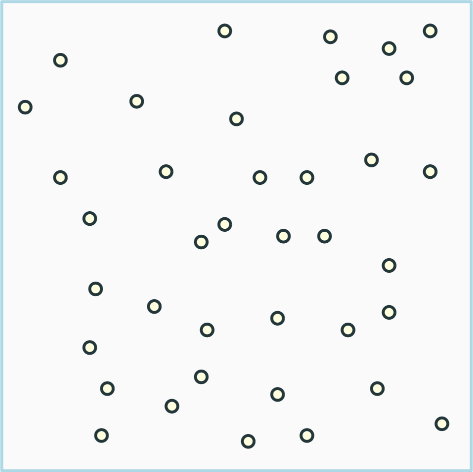

n: 26000

# Visualization

**n: 28000**

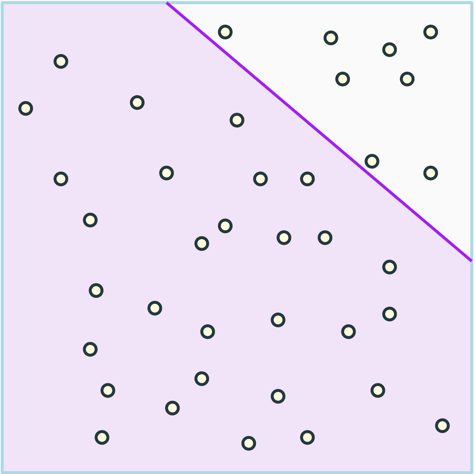
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



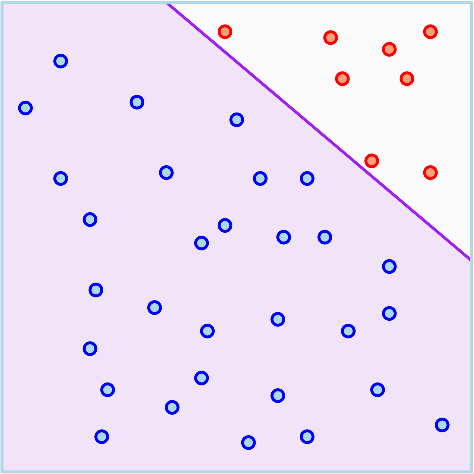
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



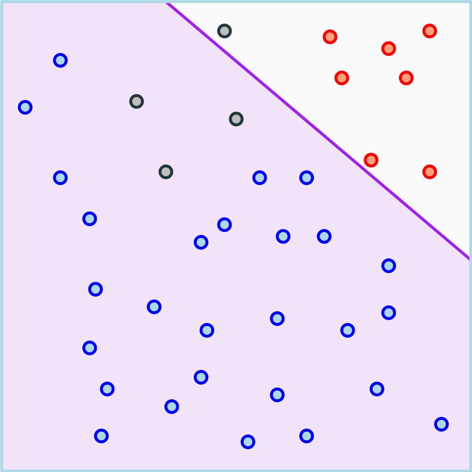
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



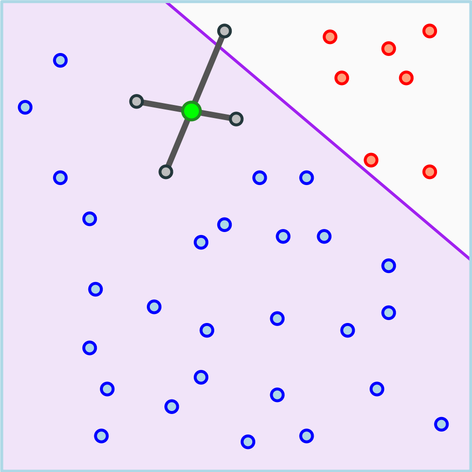
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



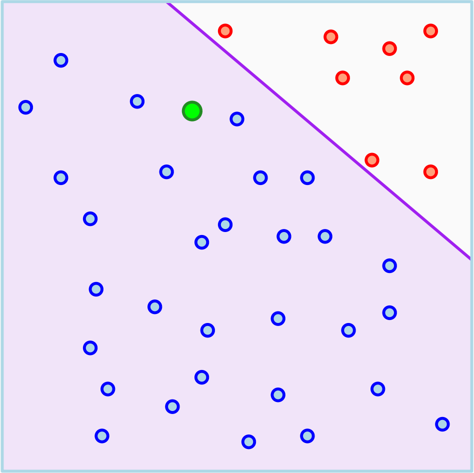
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



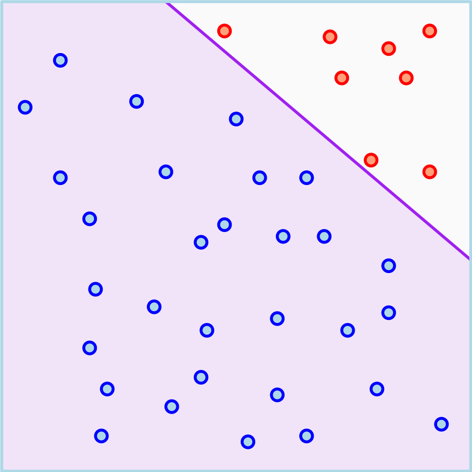
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



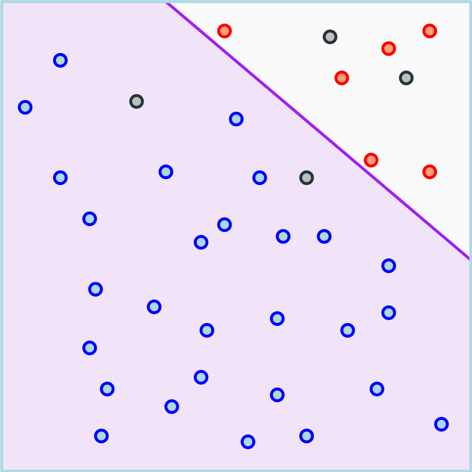
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



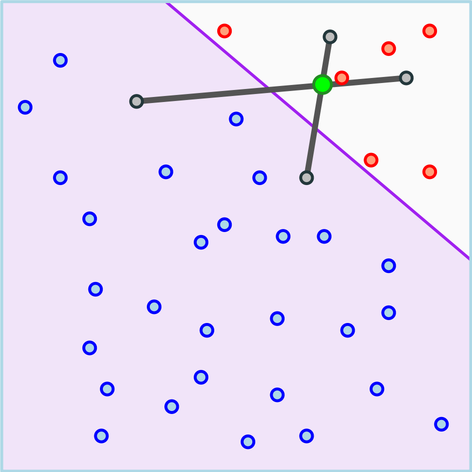
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



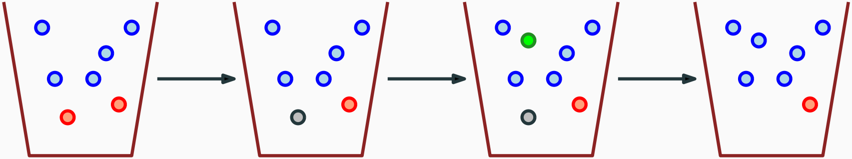
# Why does this work?

$q$ is an $\alpha$-centerpoint for $P$ $\iff$ all halfspaces $h^+$ with $|P \cap h^+| > (1-\alpha)|P|$ contain $q$



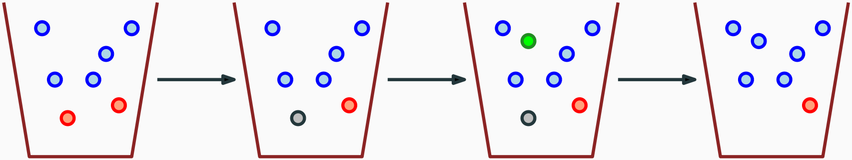
# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.



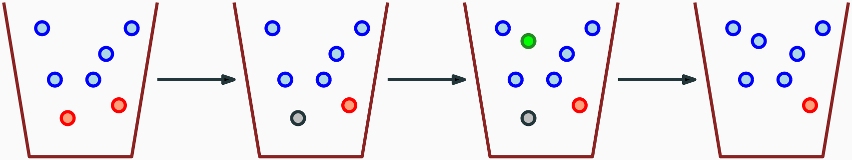
# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.
- In each round:



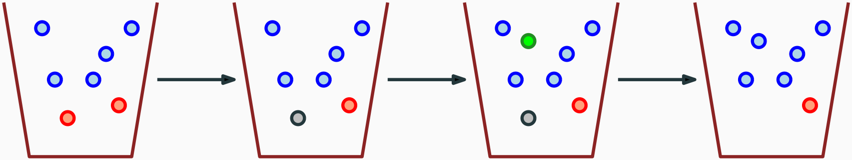
# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.
- In each round:
  1. Mark a ball for deletion.



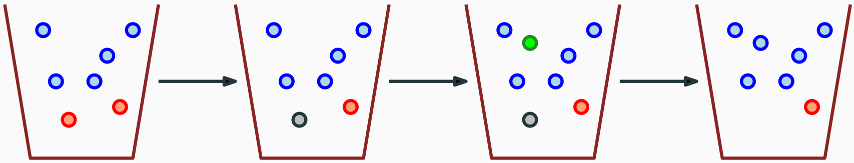
# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.
- In each round:
  1. Mark a ball for deletion.
  2. Sample $d + 2$ balls.



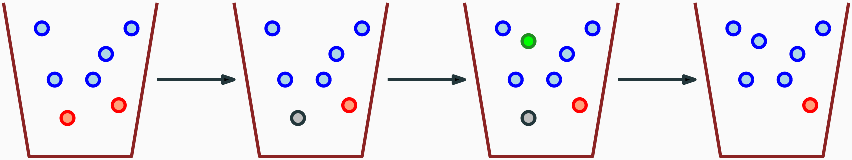
# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.
- In each round:
  1. Mark a ball for deletion.
  2. Sample $d + 2$ balls.
  3. If $\geqslant 2$ balls in sample are red, add a red ball. Otherwise, add a blue ball.

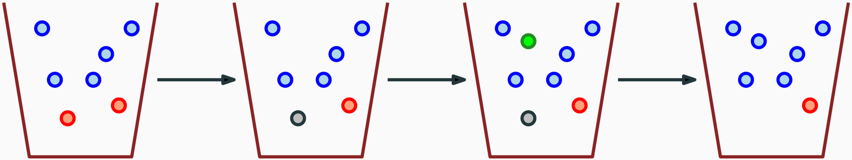


# Radon's urn

- Urn with $b$ blue balls, $r = m - b$ red balls.
- In each round:
  1. Mark a ball for deletion.
  2. Sample $d + 2$ balls.
  3. If $\geqslant 2$ balls in sample are red, add a red ball. Otherwise, add a blue ball.
  4. Remove marked ball from urn.

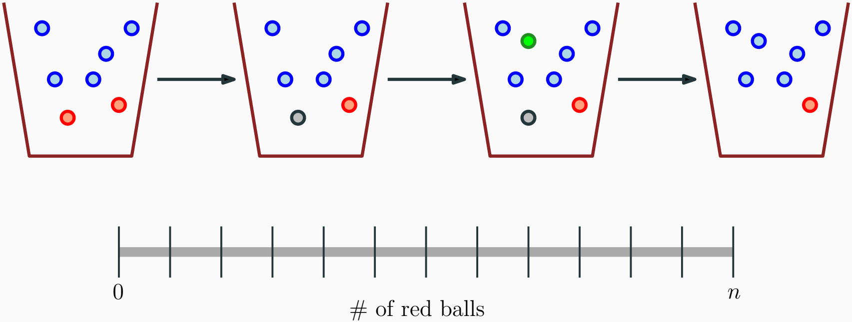


Random walk process

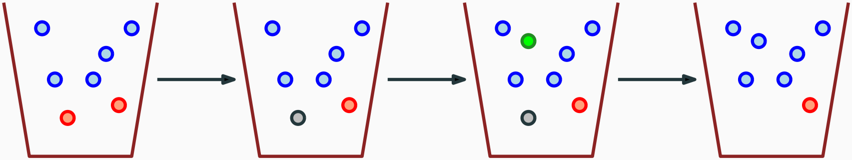


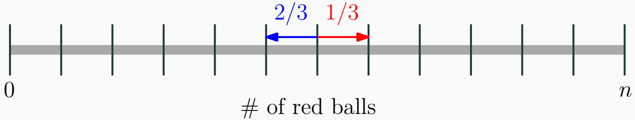
# Radon’s urn

Random walk process



Random walk process

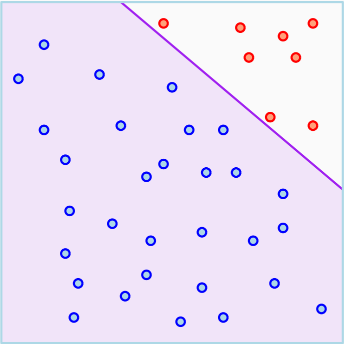




# Radon’s urn

**Problem**

Number of rounds until all balls are blue?

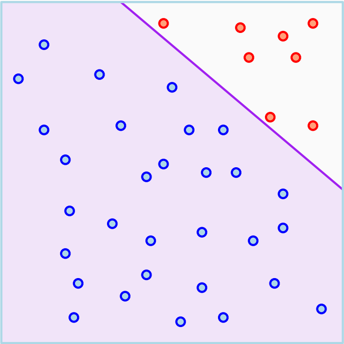


### Problem

Number of rounds until all balls are blue?

### Our result

When # of balls $m$ is sufficiently large: $O(m \log^2 m)$ rounds.



# Radon's urn

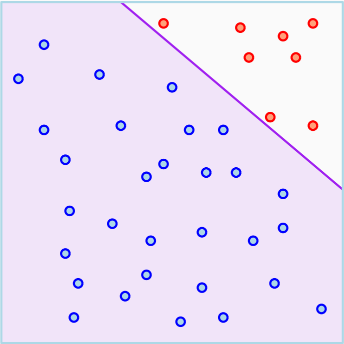
### Problem

Number of rounds until all balls are blue?

### Our result

When # of balls $m$ is sufficiently large: $O(m \log^2 m)$ rounds.

- Simulate random walk process in parallel for all $O(n^d)$ halfspaces.



### Our result

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

### Our result

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

- Radon points: quick to compute, good centerpoints

# Result

## Our result

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

- Radon points: quick to compute, good centerpoints
- Algorithm is many parallel random walks

**Our result**

$P \subset \mathbb{R}^d$: set of $n$ points.

With random sampling, $\approx 1/(d+2)^2$-centerpoint in time $O(d^7 \log d)$.

- Radon points: quick to compute, good centerpoints
- Algorithm is many parallel random walks

**Problem**

$\approx 1/d$-centerpoints in $O(\text{poly}(d))$ time?

### Definition: center nets

$P \subset \mathbb{R}^d$: set of $n$ points.

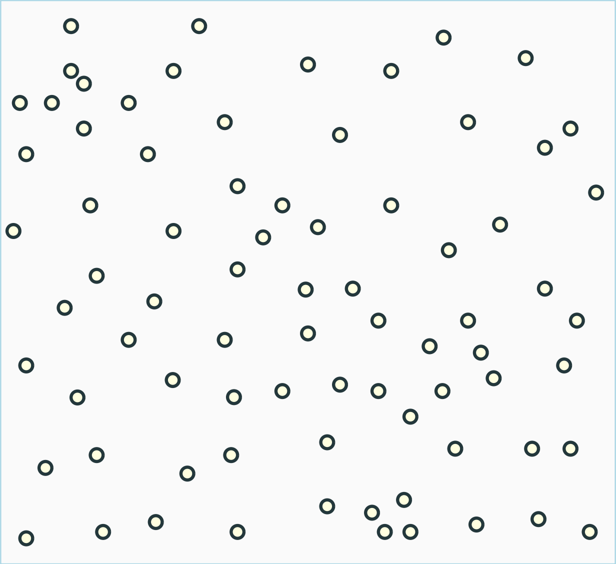
$Q \subset \mathbb{R}^d$, $(\varepsilon, \alpha)$-center net if $\forall$ convex bodies $C \subseteq \mathbb{R}^d$:

$|P \cap C| \geqslant \varepsilon n \implies \exists q \in Q \cap C$, $q$ an $\alpha$-centerpoint of $P \cap C$.

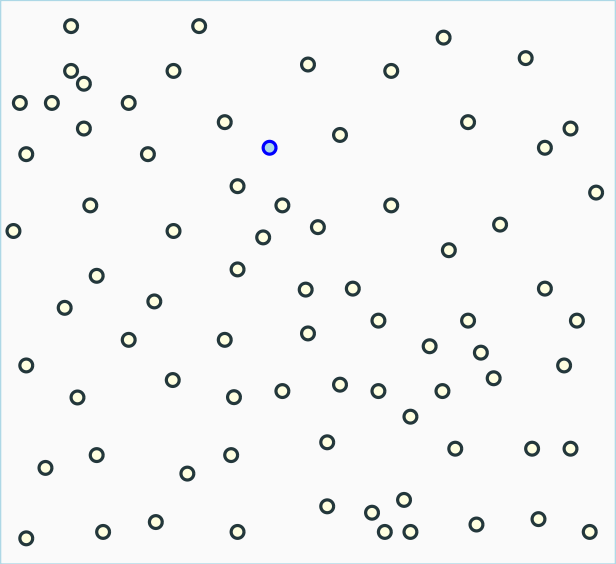
### Our result

There exists an $\left(\varepsilon, \Omega\left(\frac{1}{d \log \varepsilon^{-1}}\right)\right)$-center net for $P$ of size $\widetilde{O}\left((d^2/\varepsilon)^{d^2}\right)$.

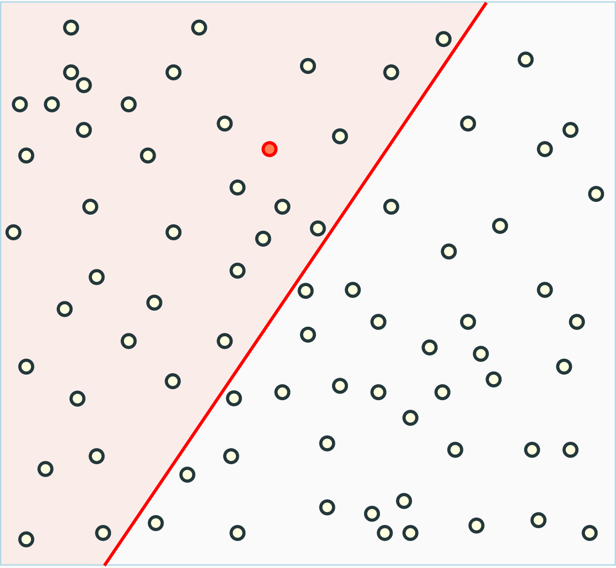
# Application: Functional nets



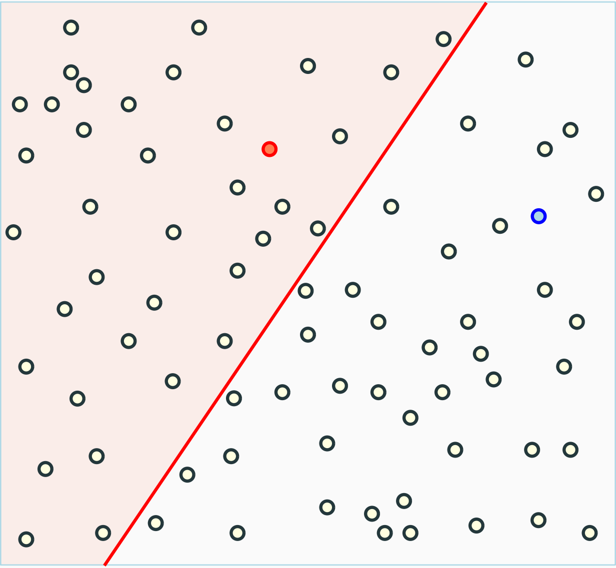
# Application: Functional nets



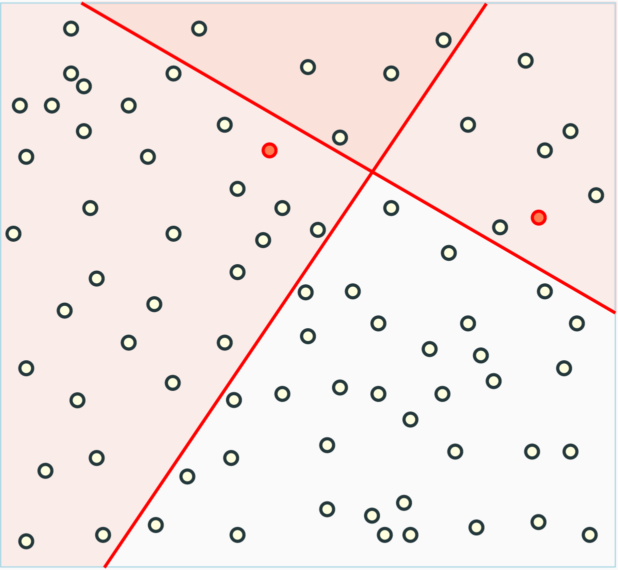
# Application: Functional nets



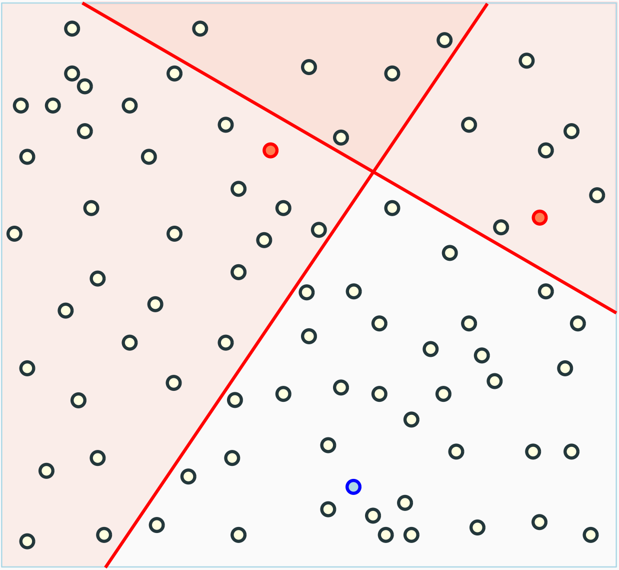
# Application: Functional nets



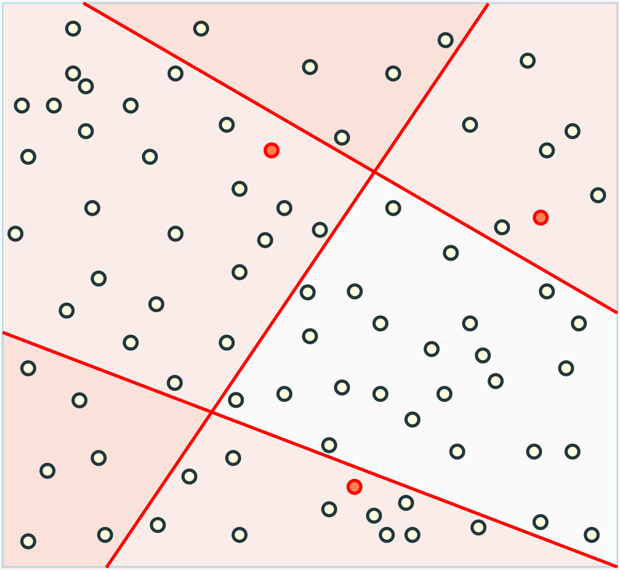
# Application: Functional nets



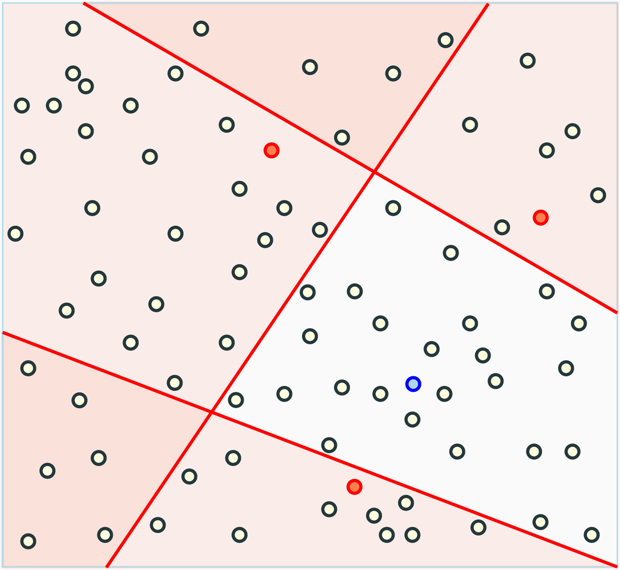
# Application: Functional nets



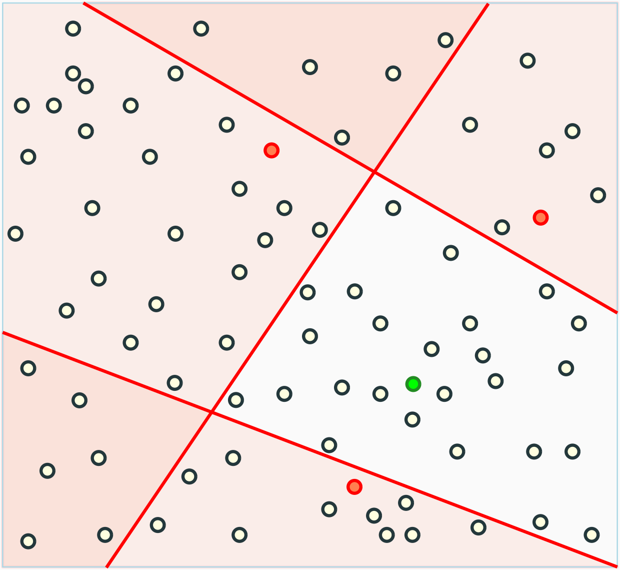
# Application: Functional nets



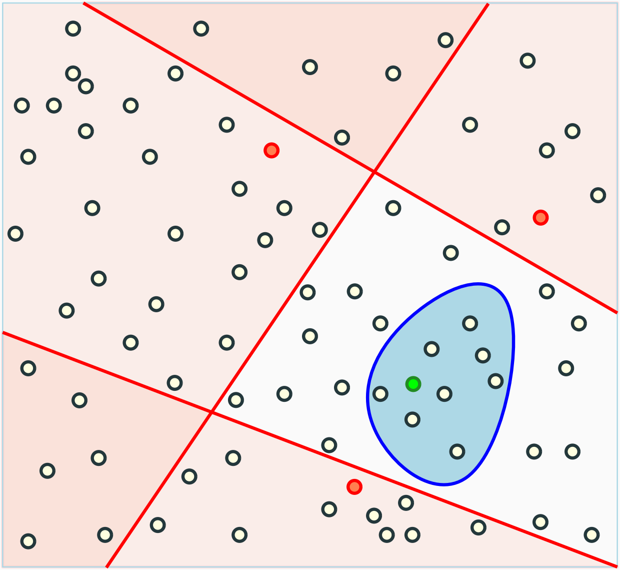
# Application: Functional nets



# Application: Functional nets



# Application: Functional nets



# Application: Functional nets

**Our result**

Can verify if $|C \cap P| \leqslant \varepsilon n$ with $O(d^2 \log \varepsilon^{-1})$ oracle queries to $C$, in $\widetilde{O}(d^9/\varepsilon)$ randomized time.

# Application: Functional nets

**Our result**

Can verify if $|C \cap P| \leqslant \varepsilon n$ with $O(d^2 \log \varepsilon^{-1})$ oracle queries to $C$, in $\widetilde{O}(d^9/\varepsilon)$ randomized time.

- Weak $\varepsilon$-nets in a different model

**Our result**

Can verify if $|C \cap P| \leqslant \varepsilon n$ with $O(d^2 \log \varepsilon^{-1})$ oracle queries to $C$, in $\widetilde{O}(d^9/\varepsilon)$ randomized time.

- Weak $\varepsilon$-nets in a different model
- Weak $\varepsilon$-nets have exponential dependency on $d$ [Matoušek and Wagner, 2004] [Mustafa and Ray, 2008]

**Our result**

Can verify if $|C \cap P| \leqslant \varepsilon n$ with $O(d^2 \log \varepsilon^{-1})$ oracle queries to $C$, in $\widetilde{O}(d^9/\varepsilon)$ randomized time.

- Weak $\varepsilon$-nets in a different model
- Weak $\varepsilon$-nets have exponential dependency on $d$ [Matoušek and Wagner, 2004] [Mustafa and Ray, 2008]
- What models can we obtain similar results with better dependency on $d$?

- T. M. Chan. *An optimal randomized algorithm for maximum Tukey depth.* 430–436, 2004.
- K. L. Clarkson, D. Eppstein, G. L. Miller, C. Sturtivant, and S.-H. Teng. *Approximating center points with iterative Radon points.* *Internat. J. Comput. Geom. Appl.*, 6: 357–377, 1996.
- G. L. Miller and D. R. Sheehy. *Approximate centerpoints with proofs.* *Comput. Geom.*, 43(8): 647–654, 2010.
- Y. Li, P. M. Long, and A. Srinivasan. *Improved bounds on the sample complexity of learning.* *J. Comput. Syst. Sci.*, 62(3): 516–527, 2001.
- J. Matoušek and U. Wagner. *New constructions of weak epsilon-nets.* *Discrete Comput. Geom.*, 32(2): 195–206, 2004.

N. H. Mustafa and S. Ray. *Weak $\varepsilon$-nets have basis of size $O(\varepsilon^{-1} \log \varepsilon^{-1})$ in any dimension.* *Comput. Geom. Theory Appl.*, 40(1): 84–91, 2008.

V Vapnik and A Chervonenkis. *On the uniform convergence of relative frequencies of events to their probabilities.* 1971.