The Hardest Halfspace

Alexander Sherstov
UCLA

Approximation by polynomials

$f : \{0,1\}^n \to \{-1,+1\}$

Approximation by polynomials

$f : \{0,1\}^n \to \{-1,+1\}$

DEFINITION.
$E(f, d)$ is the minimum error to which f can be approximated in l_∞ norm by a polynomial $p \in \mathbf{R}[x_1, x_2, \ldots, x_n]$ of degree $\leq d$:

$$E(f,d) = \min_{p:\deg p \leq d} \|f - p\|_\infty$$

Approximation by polynomials

$f : \{0,1\}^n \to \{-1,+1\}$

DEFINITION.
$E(f, d)$ is the minimum error to which f can be approximated in l_∞ norm by a polynomial $p \in \mathbf{R}[x_1, x_2, \ldots, x_n]$ of degree $\leq d$:

$$E(f,d) = \min_{p:\deg p \leq d} \|f - p\|_\infty$$

$E(f,n) = 0$

Approximation by polynomials

$f : \{0,1\}^n \to \{-1,+1\}$

DEFINITION.
$E(f, d)$ is the minimum error to which f can be approximated in l_∞ norm by a polynomial $p \in \mathbf{R}[x_1, x_2, \ldots, x_n]$ of degree $\leq d$:

$$E(f,d) = \min_{p:\deg p \leq d} \|f - p\|_\infty$$

$E(f,n) = 0$

$E(f,d) \in [0,1]$

Motivation

- **Circuit complexity**
  [PS94, SRK94, BRS95, ABFR94, KP97, KP98, S09, BH12]
- **Quantum query complexity**
  [BBC+01, BCWZ99, AS04, A05, A05, KŠW07, BKT17]
- **Communication complexity**
  [BW01, R02, BVW07, S09, S11, RS10, LS09, CA08, S08, BH12, S14, S16]
- **Learning theory**
  [TT99, KS04, KOS04, KKMS08, OS10, ACR+10]
- **Algorithm design**
  [LN90, KLS96, S09]
- **Differential privacy**
  [TUV12, CTUW14]

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \operatorname{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \operatorname{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \operatorname{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \operatorname{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \operatorname{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)
$E(h, \varepsilon n) \geq 1/3$	Paturi (1992)
$E(h, d) \geq 1 - 2^{-\Omega(n/d^2)}$	Beigel (1994)

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \mathrm{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)
$E(h, \varepsilon n) \geq 1/3$	Paturi (1992)
$E(h, d) \geq 1 - 2^{-\Omega(n/d^2)}$	Beigel (1994)
$E(h, \varepsilon n) \geq 1 - 2^{-\Omega(n)}$	[S. 2010]

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \mathrm{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)
$E(h, \varepsilon n) \geq 1/3$	Paturi (1992)
$E(h, d) \geq 1 - 2^{-\Omega(n/d^2)}$	Beigel (1994)
$E(h, \varepsilon n) \geq 1 - 2^{-\Omega(n)}$	[S. 2010]

Our problem

What is the worst-case $E(h, d)$ for a halfspace $h : \{0,1\}^n \to \{-1,+1\}$, $h(x) = \mathrm{sgn}(\sum_{i=1}^{n} z_i x_i - \theta)$?

Bound	Reference
$E(h, 1) \leq 1 - 2^{-O(n \log n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n)}$	Folklore
$E(h, 1) \geq 1 - 2^{-\Omega(n \log n)}$	Håstad (1994)
$E(h, \varepsilon n) \geq 1/3$	Paturi (1992)
$E(h, d) \geq 1 - 2^{-\Omega(n/d^2)}$	Beigel (1994)
$E(h, \varepsilon n) \geq 1 - 2^{-\Omega(n)}$, **NOT EXPLICIT**	[S. 2010]

Main result

MAIN RESULT.

There is an explicitly given halfspace $h : \{0,1\}^n \to \{-1,+1\}$ such that:

$$E(h, d) \geq 1 - \exp(-\Omega(n)) \qquad \text{for } d = 0, 1, \ldots, \varepsilon n$$

$$R(h, d) \geq 1 - \exp(-\Omega(n / d)) \qquad \text{for } d = 0, 1, \ldots, \varepsilon n$$

approximation by rational functions

Main result

MAIN RESULT.

There is an explicitly given halfspace $h : \{0,1\}^n \to \{-1,+1\}$ such that:

$$E(h, d) \geq 1 - \exp(-\Omega(n)) \qquad \text{for } d = 0, 1, \ldots, \varepsilon n$$

$$R(h, d) \geq 1 - \exp(-\Omega(n / d)) \qquad \text{for } d = 0, 1, \ldots, \varepsilon n$$

approximation by rational functions

- Optimal for both E and R
- Quadratic improvement on previous explicit constructions

Main result

COROLLARY.

There are explicitly given halfspaces $h_1, h_2 : \{0,1\}^n \to \{-1,+1\}$ such that:

$$E(\, h_1 \wedge h_2, \Omega(n)) \;= 1$$
$$R(\, h_1 \wedge h_2, \Omega(n)) \;= 1$$

Main result

COROLLARY.

There are explicitly given halfspaces $h_1, h_2 : \{0,1\}^n \to \{-1,+1\}$ such that:

$$E(h_1 \wedge h_2, \Omega(n)) = 1$$
$$R(h_1 \wedge h_2, \Omega(n)) = 1$$

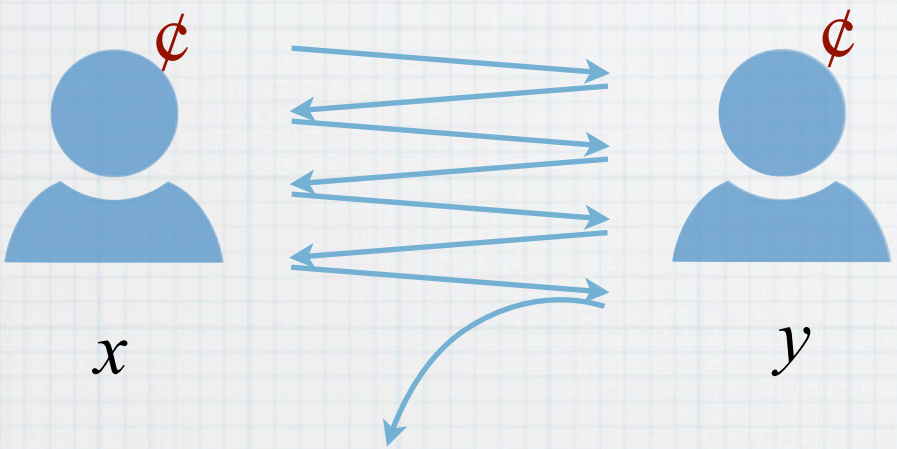
- Optimal for both E and R
- Quadratic improvement on previous explicit constructions

Application 1:

Unbounded-error communication

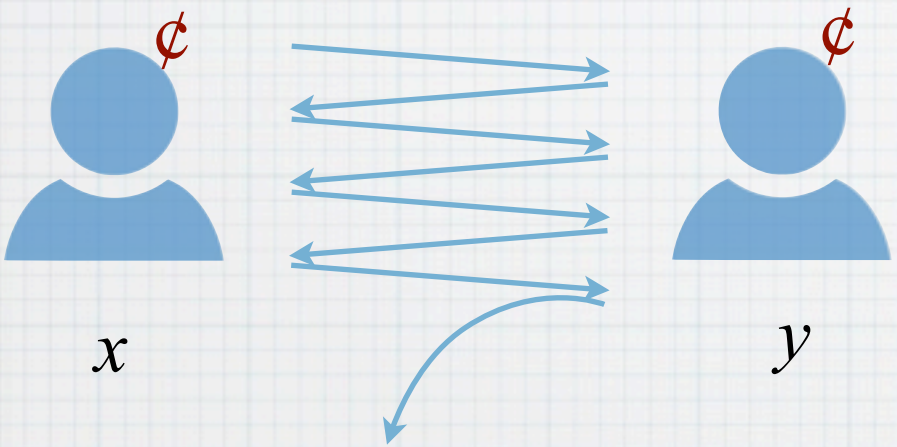
UPP versus PP

$f : \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$



UPP versus PP

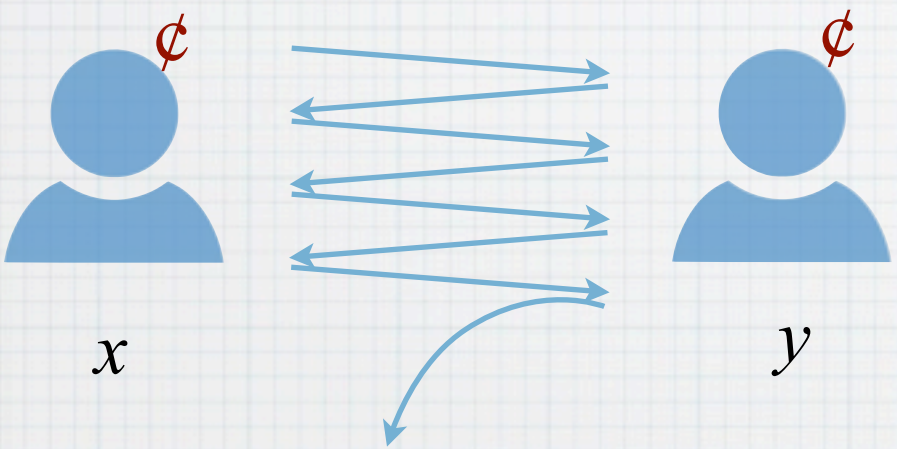
$f : \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$



$R_\epsilon(f) =$ minimum cost of an ε-error protocol for f

UPP versus PP

$f : \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$

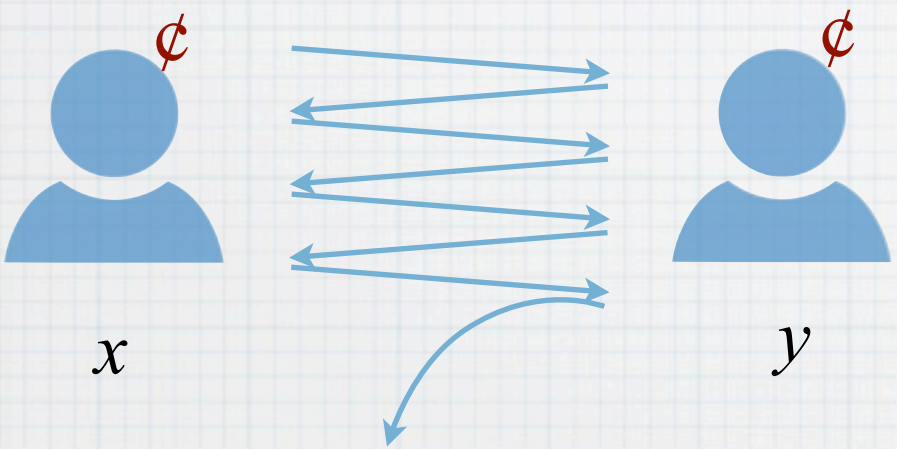


$R_\epsilon(f) =$ minimum cost of an ε-error protocol for f

$$\mathrm{UPP}(f) = \min_{0<\epsilon<1/2} R_\epsilon(f)$$

UPP versus PP

$f : \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$



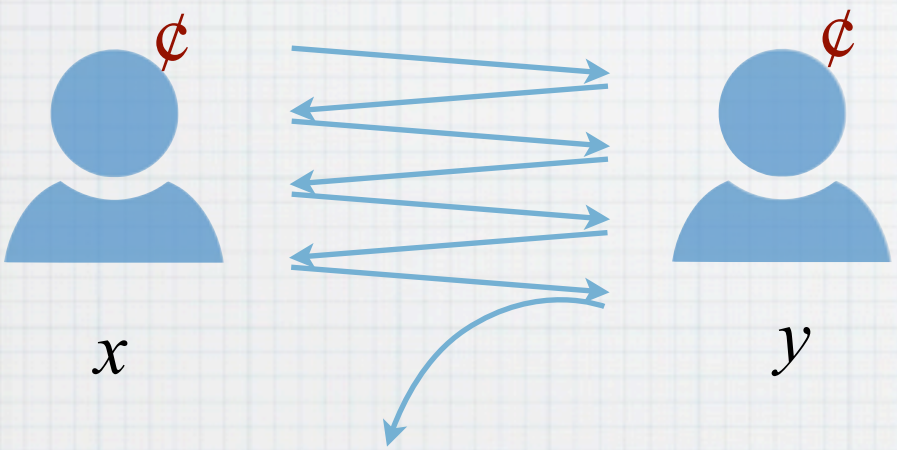
$R_\epsilon(f) =$ minimum cost of an ε-error protocol for f

$$\mathrm{UPP}(f) = \min_{0<\epsilon<1/2} R_\epsilon(f)$$

$$\mathrm{PP}(f) = \min_{0<\epsilon<1/2} \left\{ R_\epsilon(f) + \log \frac{1}{\frac{1}{2} - \epsilon} \right\}$$

UPP versus PP

$f : \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$



$R_\epsilon(f) =$ minimum cost of an ε-error protocol for f

$$\mathrm{UPP}(f) = \min_{0<\epsilon<1/2} R_\epsilon(f)$$

a.k.a. sign-rank

$$\mathrm{PP}(f) = \min_{0<\epsilon<1/2} \left\{ R_\epsilon(f) + \log \frac{1}{\frac{1}{2} - \epsilon} \right\}$$

a.k.a. discrepancy

UPP versus PP

$$\mathsf{UPP} = \{\{f_n\}_{n=1}^{\infty} : \mathrm{UPP}(f_n) = \log^{O(1)} n\}$$

$$\mathsf{PP} = \{\{f_n\}_{n=1}^{\infty} : \mathrm{PP}(f_n) = \log^{O(1)} n\}$$

UPP versus PP

$$\mathsf{UPP} = \{\{f_n\}_{n=1}^{\infty} : \mathrm{UPP}(f_n) = \log^{O(1)} n\}$$

$$\mathsf{PP} = \{\{f_n\}_{n=1}^{\infty} : \mathrm{PP}(f_n) = \log^{O(1)} n\}$$

Babai, Frankl, & Simon (1986): $\mathsf{PP} \subsetneq \mathsf{UPP}$?

UPP versus PP

Babai, Frankl, & Simon (1986): $\mathsf{PP} \subsetneq \mathsf{UPP}$?

UPP versus PP

Babai, Frankl, & Simon (1986): PP $\subsetneq$ UPP?

UPP vs. PP	Reference
$O(\log n)$ vs. $\Omega(n^{1/3})$	[Buhrman et al., 2007]
$O(\log n)$ vs. $\Omega(n^{1/2})$	[S. 2007]

UPP versus PP

Babai, Frankl, & Simon (1986): PP $\subsetneq$ UPP?

UPP vs. PP	Reference
$O(\log n)$ vs. $\Omega(n^{1/3})$	[Buhrman et al., 2007]
$O(\log n)$ vs. $\Omega(n^{1/2})$	[S. 2007]
$O(\log n)$ vs. $\Omega(n^{2/5})$	[Thaler, 2016]

UPP versus PP

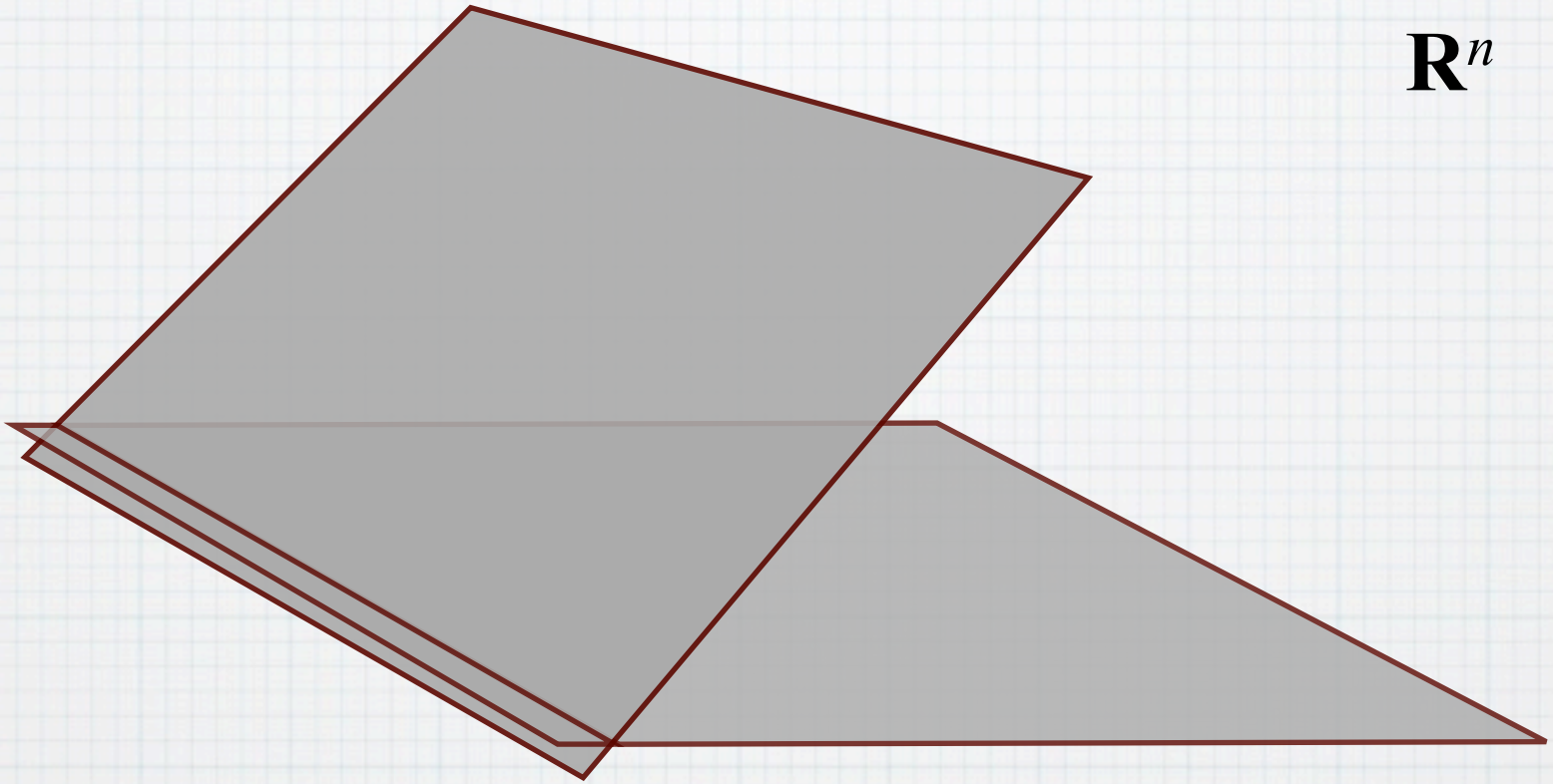
Babai, Frankl, & Simon (1986): PP $\subsetneq$ UPP?

UPP vs. PP	Reference
$O(\log n)$ vs. $\Omega(n^{1/3})$	[Buhrman et al., 2007]
$O(\log n)$ vs. $\Omega(n^{1/2})$	[S. 2007]
$O(\log n)$ vs. $\Omega(n^{2/5})$	[Thaler, 2016]
$O(\log n)$ vs. $\Omega(n)$	This work

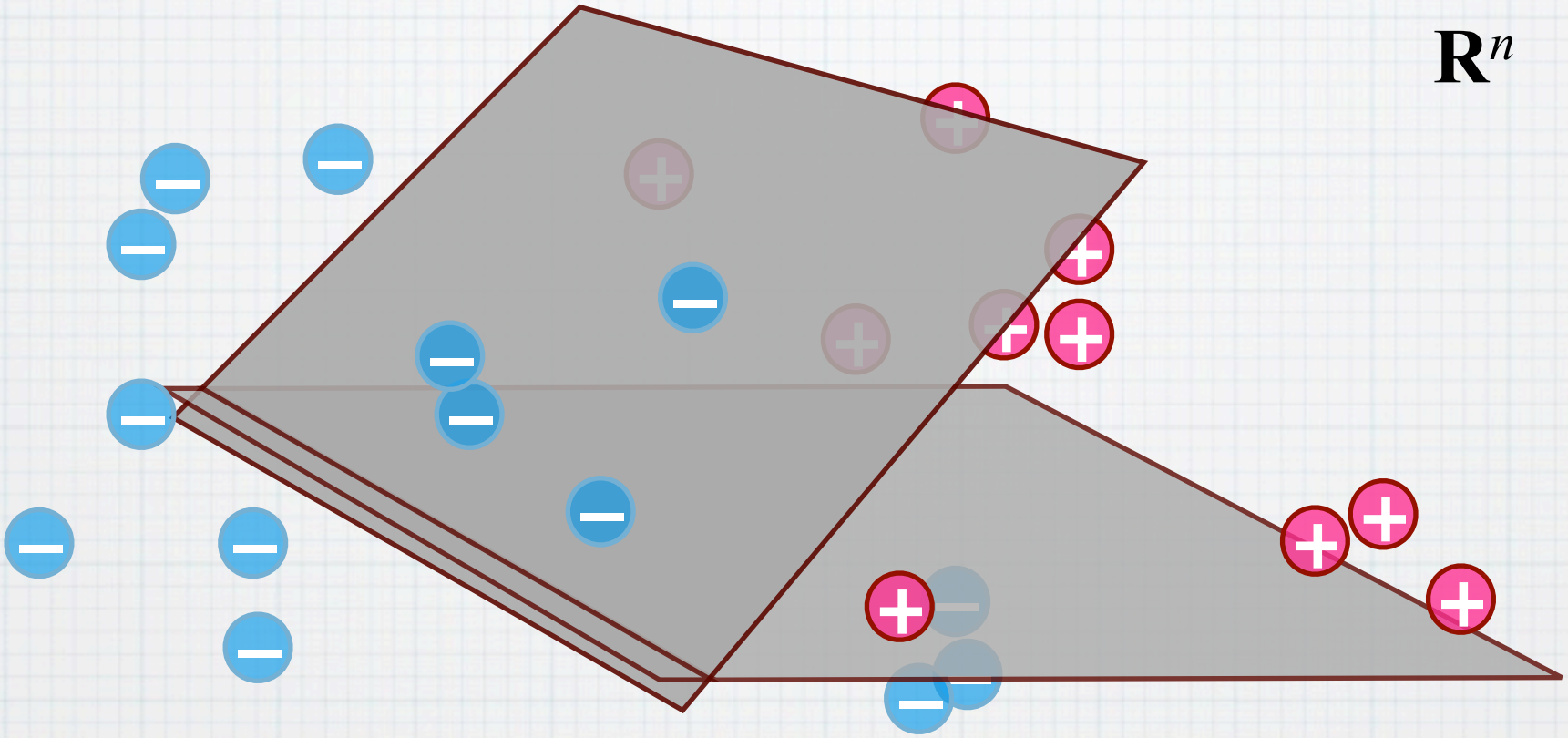
Application 2:

Learning intersections of halfspaces

The problem

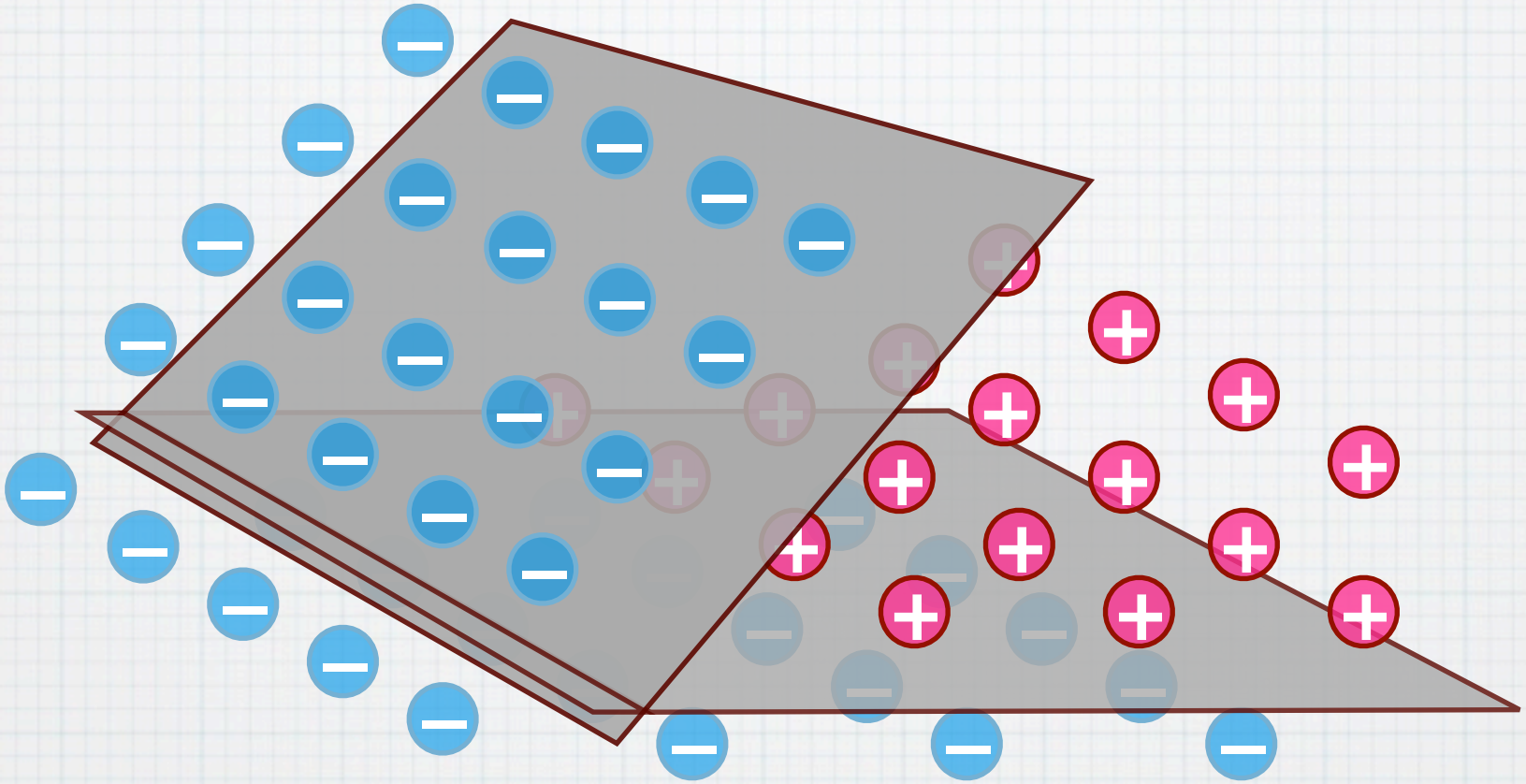


The problem



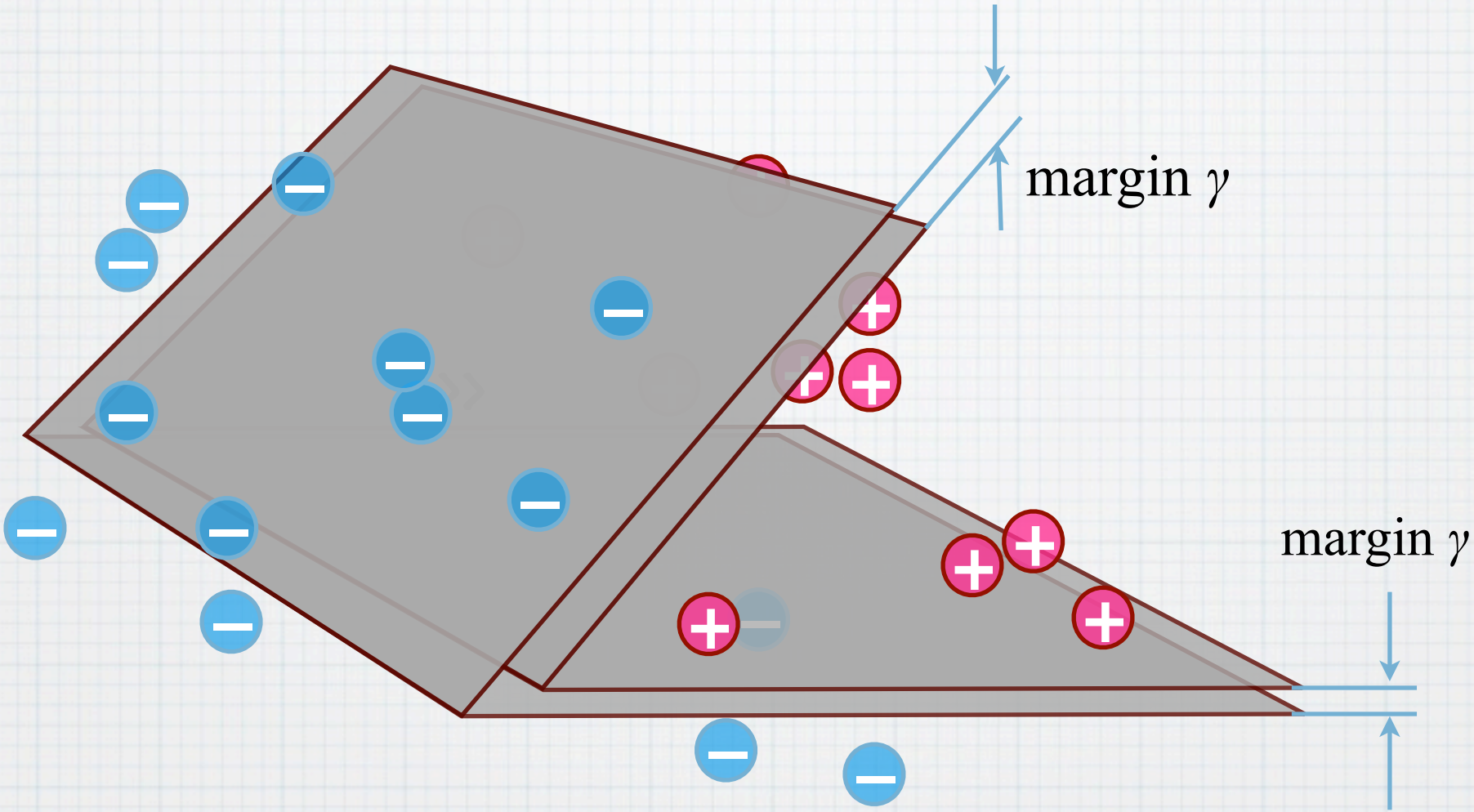
Learn intersection from examples.

Restricted distributions



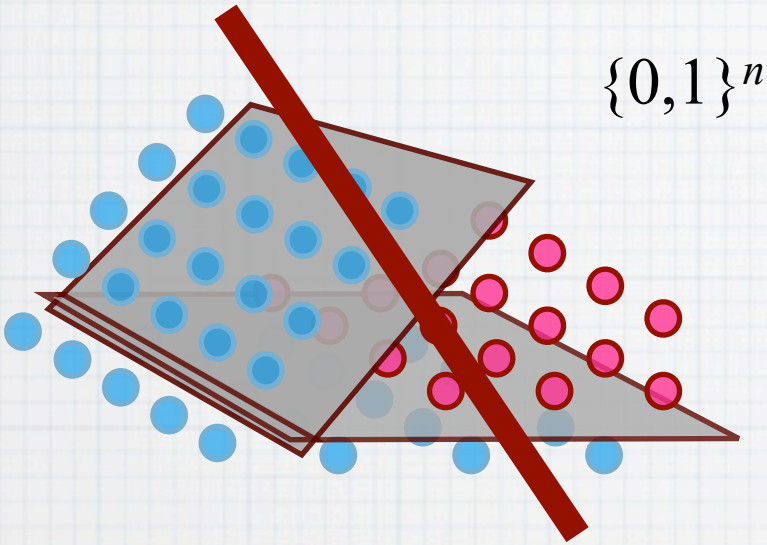
Symmetric on $\mathbf{R}^n$	[Baum 90]
Uniform on $\mathbf{S}^{n-1}$	[Blum & Kannan '93, Vempala, '97]
Uniform on $\{0,1\}^n$	[Klivans, O'Donnell, & Servedio, '02]
Log-concave on $\mathbf{R}^n$	[Klivans, Long, & Tang, '09]

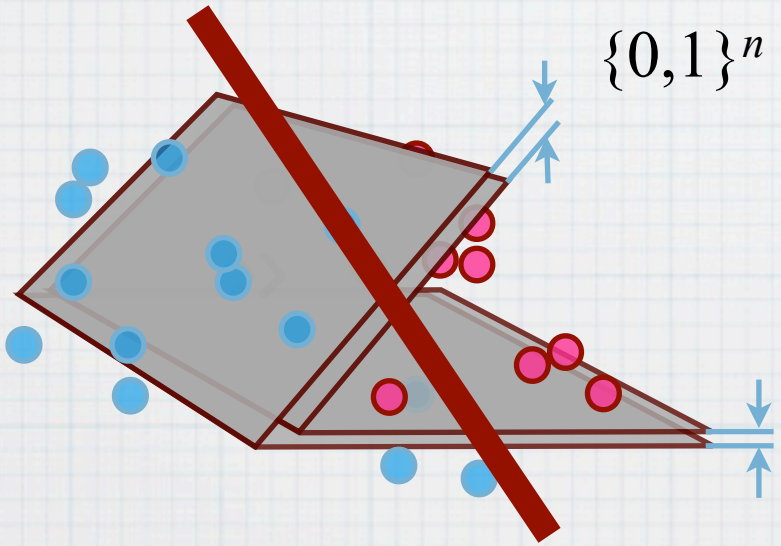
Large margin



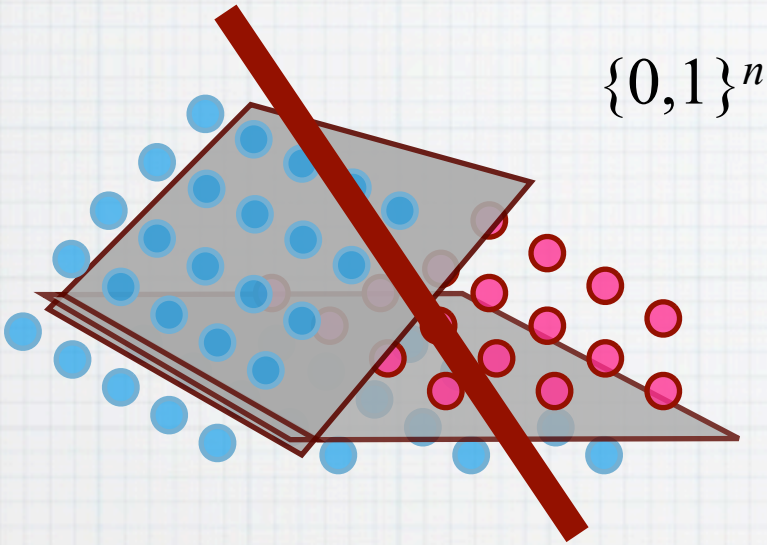
Random projection [Arriaga & Vempala '99]
[Klivans & Servedio '04]

What about general case?

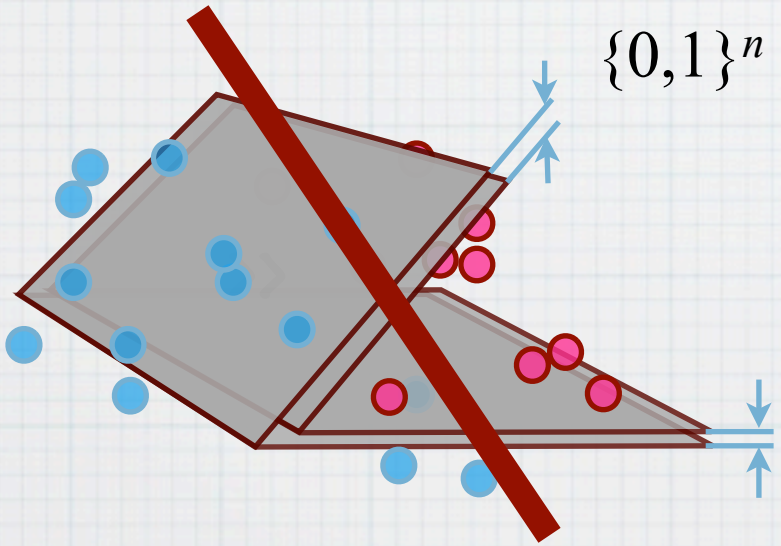




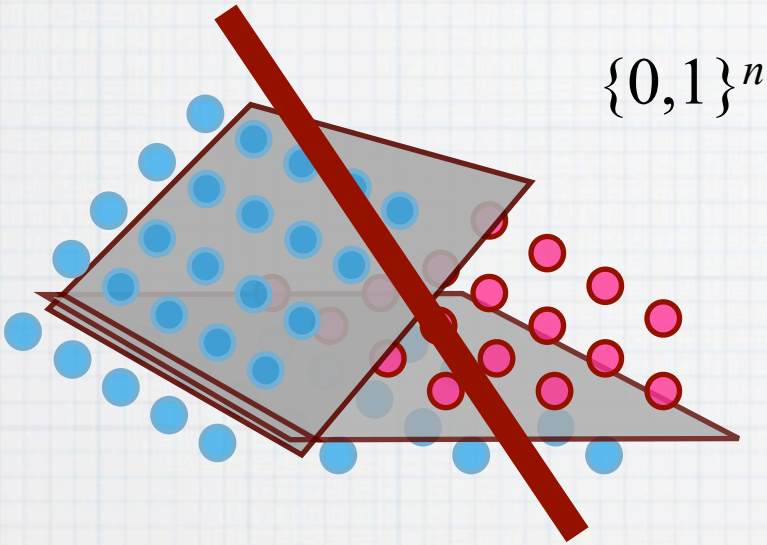
What about general case?

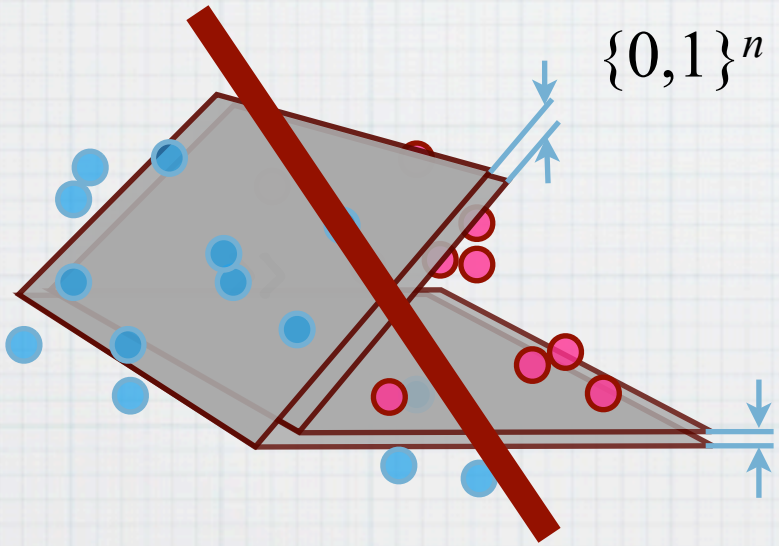


No algorithm better than $2^{\Theta(n)}$.



What about general case?



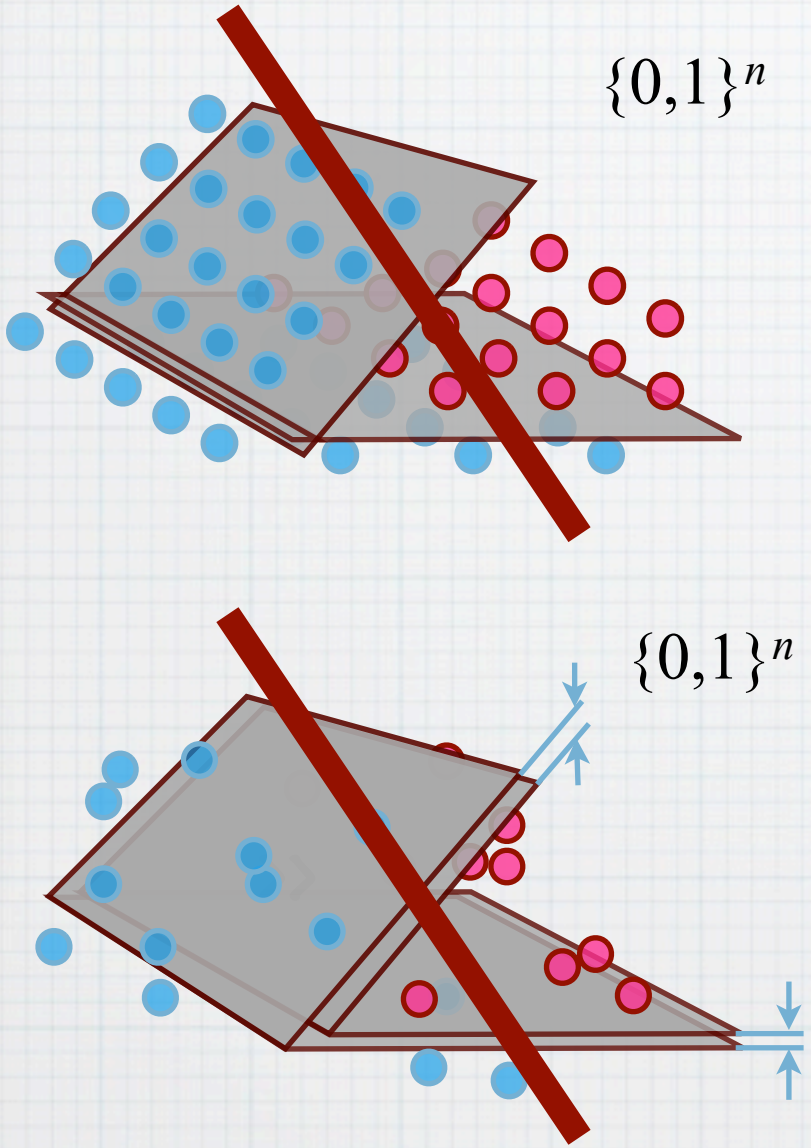


No algorithm better than $2^{\Theta(n)}$.

Hardness:

- **NP**-hardness for proper learning
  - [Blum and Rivest, '88]
  - [Alekhnovich, Braverman, Feldman, Klivans, and Pitassi, '04]
  - [Khot and Saket, '08]

What about general case?



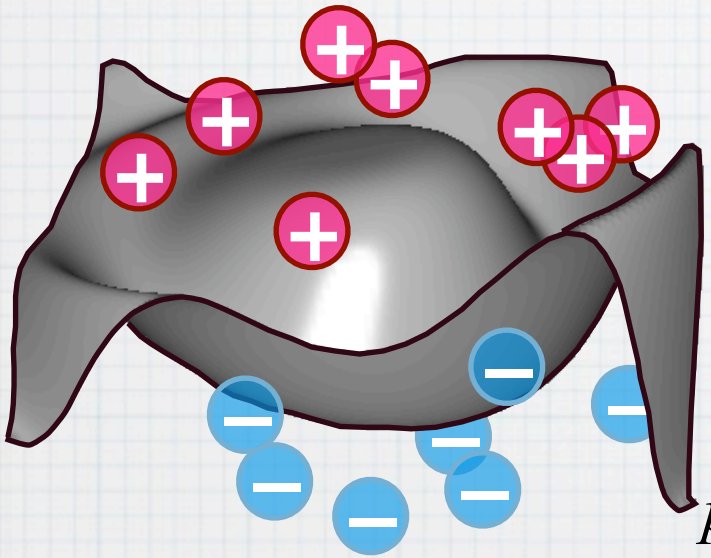
No algorithm better than $2^{\Theta(n)}$.

Hardness:

- **NP**-hardness for proper learning
  [Blum and Rivest, '88]
  [Alekhnovich, Braverman, Feldman, Klivans, and Pitassi, '04]
  [Khot and Saket, '08]
- Cryptographic hardness for n^{ε} halfspaces
  [Klivans and S., '07], using [Regev'05]
- Lower bounds for n^{ε} halfspaces in SQ model
  [Klivans and S., '06]

The polynomial method

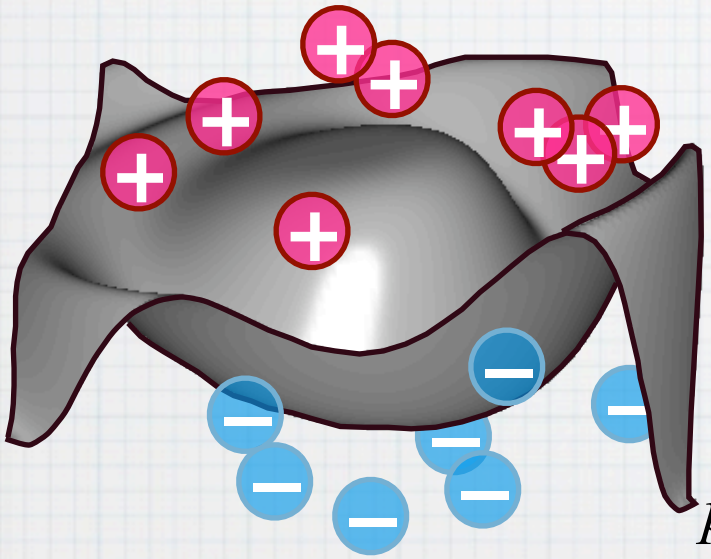
Low-degree polynomial $p : \{0,1\}^n \to \mathbf{R}$



$p(x) = 0$

The polynomial method

Low-degree polynomial $p : \{0,1\}^n \to \mathbf{R}$

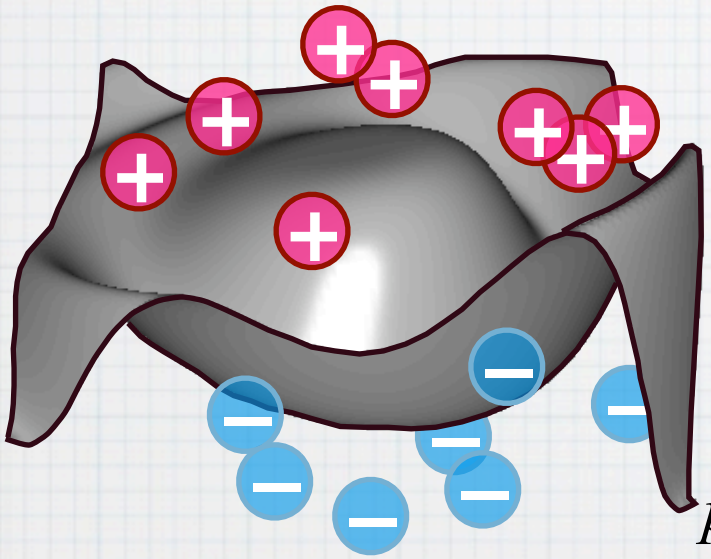


$p(x) = 0$

Can learn $f(x) = \operatorname{sgn} p(x)$
in time $n^{O(\deg p)}$ w.r.t. any distribution.

The polynomial method

Low-degree polynomial $p : \{0,1\}^n \to \mathbf{R}$



$p(x) = 0$

Can learn $f(x) = \operatorname{sgn} p(x)$
in time $n^{O(\deg p)}$ w.r.t. any distribution.

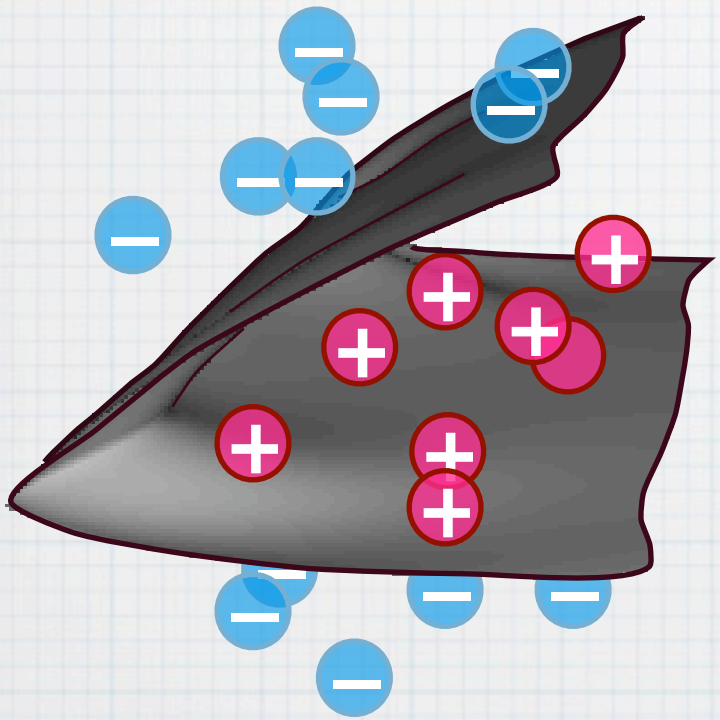
$\Rightarrow$

- $\exp(n^{1/3})$-time algorithm for DNF
  [Klivans and Servedio, 03]
- $\exp(n^{1/2})$-time algorithm for read-once formulas
  [O'Donnell and Servedio, 03]
  [Ambainis, Childs, Reichardt, Spalek, Zhang, '07]

Intersection of two halfspaces

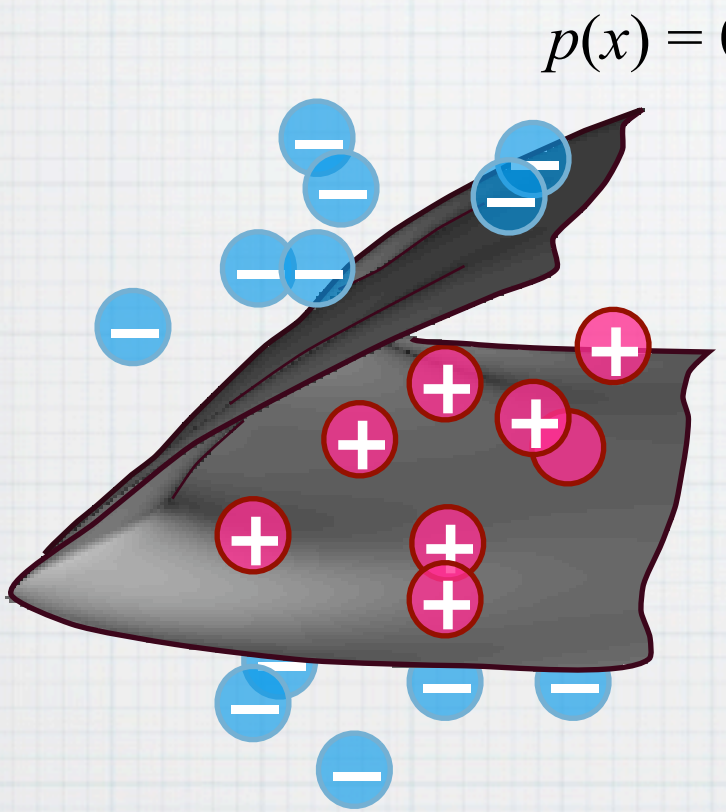
Klivans (2002): What is the worst-case threshold degree?

$p(x) = 0$



Intersection of two halfspaces

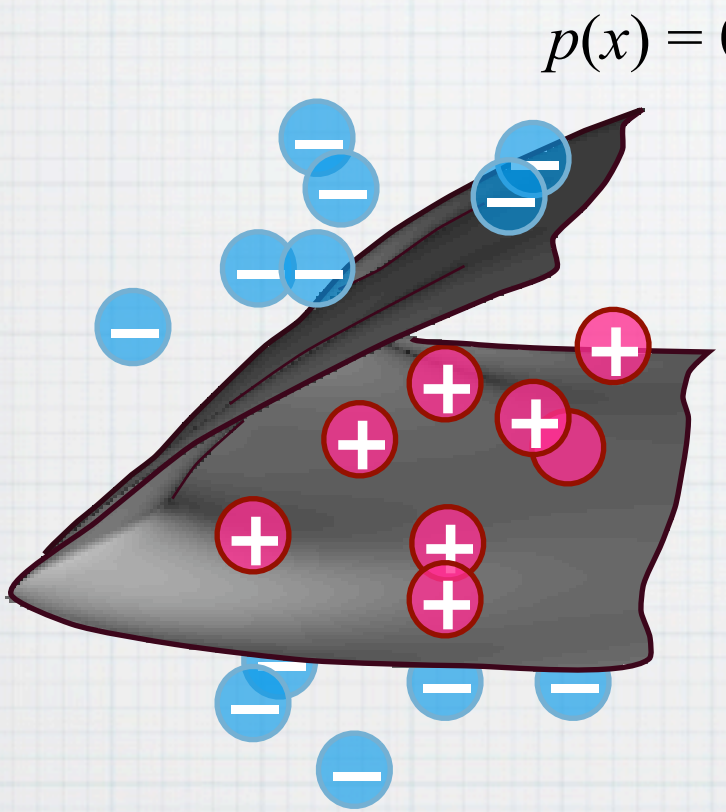
Klivans (2002): What is the worst-case threshold degree?



Bound	Reference
$\omega(1)$	[Minsky & Papert, 1969]

Intersection of two halfspaces

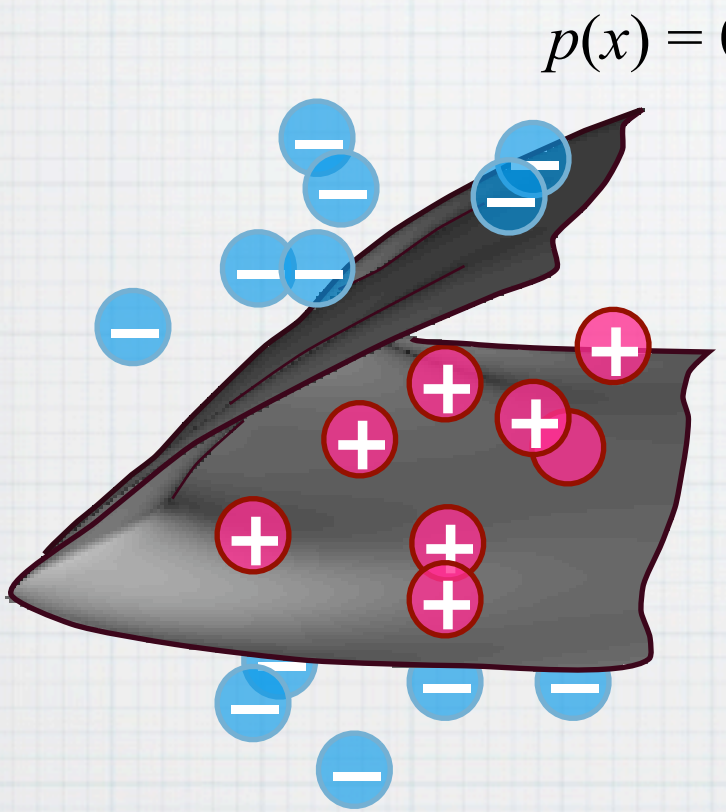
Klivans (2002): What is the worst-case threshold degree?



Bound	Reference
$\omega(1)$	[Minsky & Papert, 1969]
$\Omega\left(\frac{\log n}{\log \log n}\right)$	[O'Donnell & Servedio, 2002]

Intersection of two halfspaces

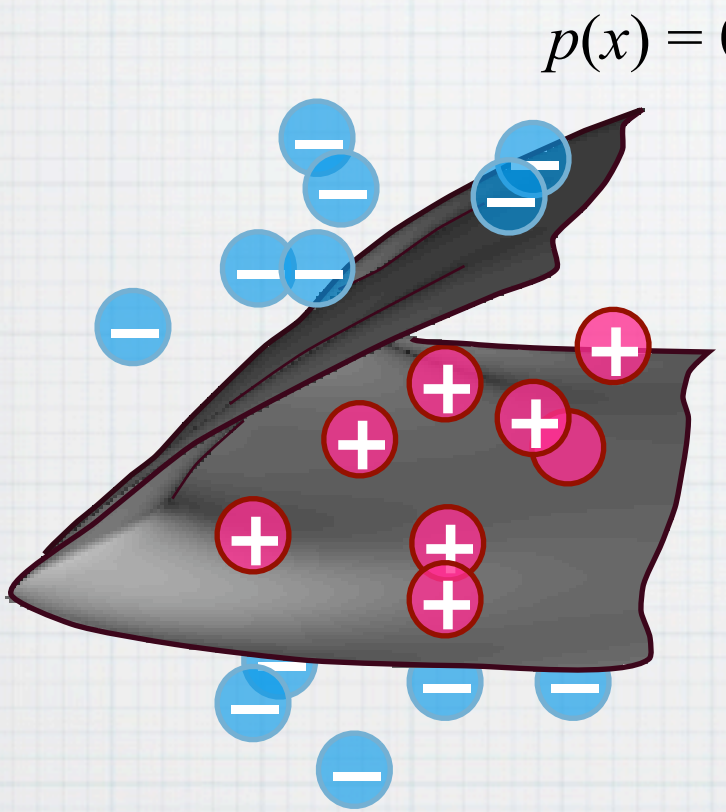
Klivans (2002): What is the worst-case threshold degree?



Bound	Reference
$\omega(1)$	[Minsky & Papert, 1969]
$\Omega\left(\frac{\log n}{\log \log n}\right)$	[O'Donnell & Servedio, 2002]
$\Omega(n^{1/2})$	[S. 2009]

Intersection of two halfspaces

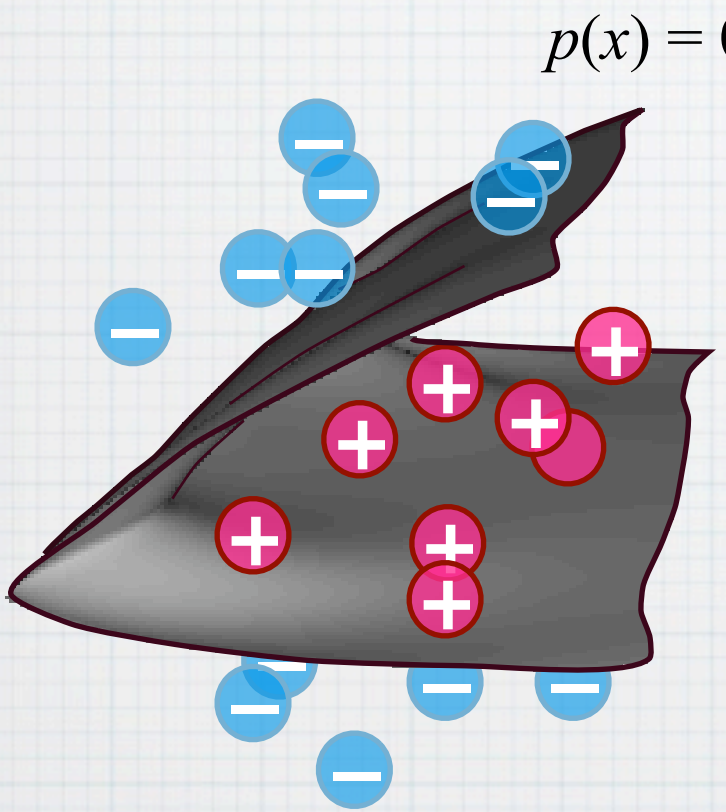
Klivans (2002): What is the worst-case threshold degree?



Bound	Reference
$\omega(1)$	[Minsky & Papert, 1969]
$\Omega\left(\frac{\log n}{\log \log n}\right)$	[O'Donnell & Servedio, 2002]
$\Omega(n^{1/2})$	[S. 2009]
$\Omega(n)$ **NOT EXPLICIT**	[S. 2010]

Intersection of two halfspaces

Klivans (2002): What is the worst-case threshold degree?



Bound	Reference
$\omega(1)$	[Minsky & Papert, 1969]
$\Omega\left(\frac{\log n}{\log \log n}\right)$	[O'Donnell & Servedio, 2002]
$\Omega(n^{1/2})$	[S. 2009]
$\Omega(n)$ **NOT EXPLICIT**	[S. 2010]
$\Omega(n)$, explicit	This work

Proof strategy

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Black-box approximants: $S(\sum z_i x_i - \theta)$ for some polynomial S or rational function S

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Black-box approximants: $S(\sum z_i x_i - \theta)$ for some polynomial S or rational function S

Optimal if $z_1 = z_2 = \cdots = z_n$ (Minsky & Papert 1968)

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Black-box approximants: $S(\sum z_i x_i - \theta)$ for some polynomial S or rational function S

Optimal if $z_1 = z_2 = \cdots = z_n$ (Minsky & Papert 1968)

Pathetic in general:

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} (-2)^i x_i + 1\right)$$

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Black-box approximants: $S(\sum z_i x_i - \theta)$ for some polynomial S or rational function S

Optimal if $z_1 = z_2 = \cdots = z_n$ (Minsky & Papert 1968)

Pathetic in general:

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} (-2)^i x_i + 1\right) \equiv \lim_{M \to \infty} \frac{\sum_{i=1}^{n} (-M)^i x_i + 1}{\sum_{i=1}^{n} M^i x_i + 1}$$

Proof strategy

Given: $f(x) = \text{sgn}(\sum z_i x_i - \theta)$ for some fixed $z_1, \ldots, z_n, \theta$

Black-box approximants: $S(\sum z_i x_i - \theta)$ for some polynomial S or rational function S

Optimal if $z_1 = z_2 = \cdots = z_n$ (Minsky & Papert 1968)

Pathetic in general:

degree 1!

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} (-2)^i x_i + 1\right) \equiv \lim_{M \to \infty} \frac{\sum_{i=1}^{n} (-M)^i x_i + 1}{\sum_{i=1}^{n} M^i x_i + 1}$$

Proof strategy

Will construct $z_1, \ldots, z_n, \theta$ such that:

- range of sum $\sum z_i x_i - \theta$ contains $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$
- for approximants of degree $\leq \varepsilon n$, weighted sum is as good as individual bits.

Proof strategy

Will construct $z_1, \ldots, z_n, \theta$ such that:

- range of sum $\sum z_i x_i - \theta$ contains $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$
- for approximants of degree $\leq \varepsilon n$, weighted sum is as good as individual bits.

$\Longrightarrow$

Any approximant of degree $\leq \varepsilon n$ for

$$f(x) = \operatorname{sgn}(\sum z_i x_i - \theta)$$

gives a univariate approximant for sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$.

Proof strategy

Will construct $z_1, \ldots, z_n, \theta$ such that:

- range of sum $\sum z_i x_i - \theta$ contains $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$
- for approximants of degree $\leq \varepsilon n$, weighted sum is as good as individual bits.

$\Longrightarrow$ Any approximant of degree $\leq \varepsilon n$ for

$$f(x) = \operatorname{sgn}(\sum z_i x_i - \theta)$$

gives a univariate approximant for sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$.

$\Longrightarrow$

$$E(f, d) \geq 1 - \exp(-\Omega(n)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n,$$
$$R(f, d) \geq 1 - \exp(-\Omega(n / d)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n.$$

Step 1:
Discrepancy mod m

Discrepancy defined

$Z = \{z_1, z_2, \ldots, z_n\}$, multiset of integers

Discrepancy defined

$Z = \{z_1, z_2, \ldots, z_n\}$, multiset of integers

DEFINITION (DISCREPANCY MOD m).

$$\mathrm{disc}_m(Z) = \max_{k=1,2,\ldots,m-1} \left| \frac{1}{n} \sum_{i=1}^{n} \omega^{kz_i} \right|,$$

where ω is any m-th root of unity.

Discrepancy defined

$Z = \{z_1, z_2, \ldots, z_n\}$, multiset of integers

DEFINITION (DISCREPANCY MOD m).

$$\operatorname{disc}_m(Z) = \max_{k=1,2,\ldots,m-1} \left| \frac{1}{n} \sum_{i=1}^{n} \omega^{kz_i} \right|,$$

where ω is any m-th root of unity.

maximum magnitude of a nonconstant Fourier coefficient of the frequency vector of Z

Discrepancy defined

$Z = \{z_1, z_2, \ldots, z_n\}$, multiset of integers

DEFINITION (DISCREPANCY MOD m).

$$\operatorname{disc}_m(Z) = \max_{k=1,2,\ldots,m-1} \left| \frac{1}{n} \sum_{i=1}^{n} \omega^{kz_i} \right|,$$

where ω is any m-th root of unity.

maximum magnitude of a nonconstant Fourier coefficient of the frequency vector of Z

- $\operatorname{disc}_m(Z) = \operatorname{disc}_m(Z \bmod m)$

Discrepancy defined

$Z = \{z_1, z_2, \ldots, z_n\}$, multiset of integers

DEFINITION (DISCREPANCY MOD m).

$$\operatorname{disc}_m(Z) = \max_{k=1,2,\ldots,m-1} \left| \frac{1}{n} \sum_{i=1}^{n} \omega^{kz_i} \right|,$$

where ω is any m-th root of unity.

maximum magnitude of a nonconstant Fourier coefficient of the frequency vector of Z

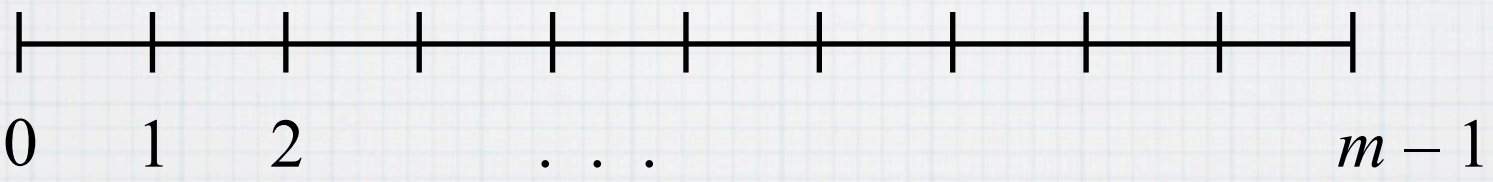
- $\operatorname{disc}_m(Z) = \operatorname{disc}_m(Z \bmod m)$
- $\operatorname{disc}_m(\{0, 1, 2, \ldots, m-1\}) = 0$

Discrepancy pictorially

Intuition: $\mathrm{disc}_m(Z)$ measures aperiodicity/balancedness

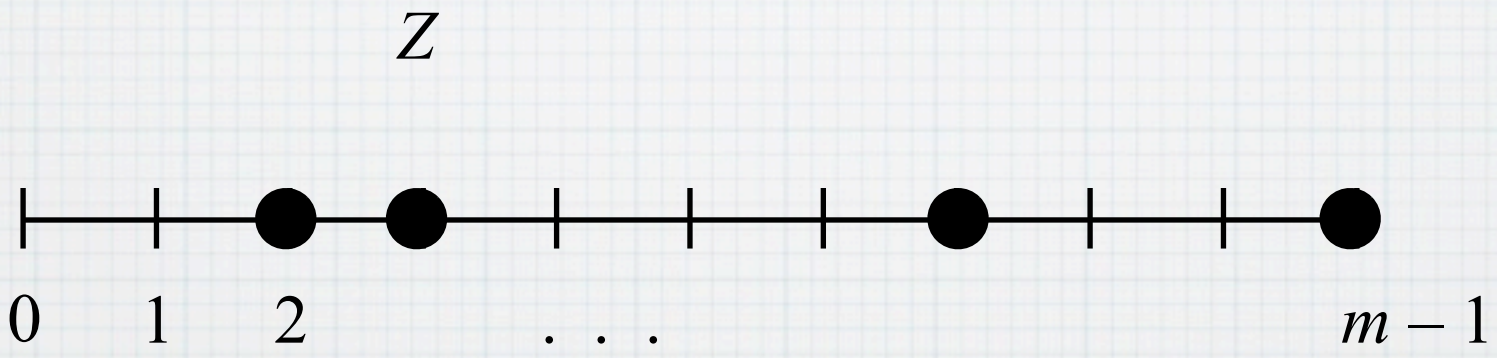
Discrepancy pictorially

Intuition: $\mathrm{disc}_m(Z)$ measures aperiodicity/balancedness



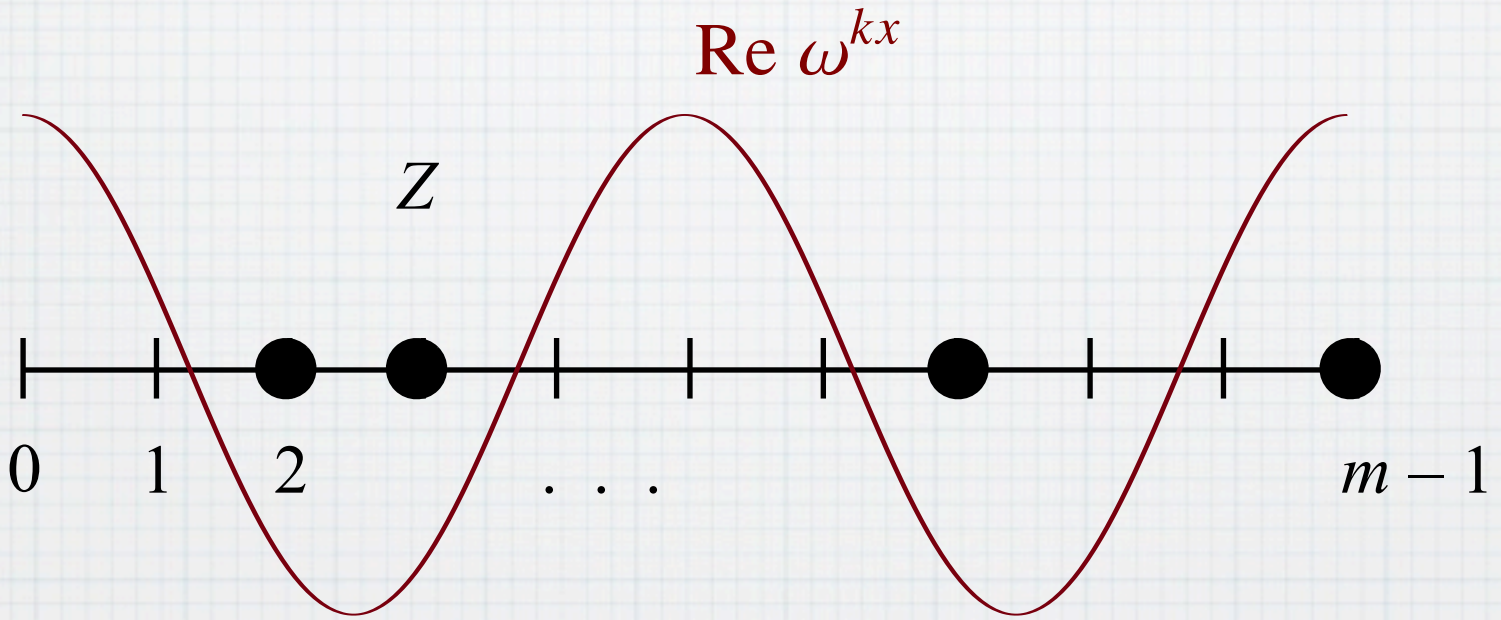
Discrepancy pictorially

Intuition: $\mathrm{disc}_m(Z)$ measures aperiodicity/balancedness



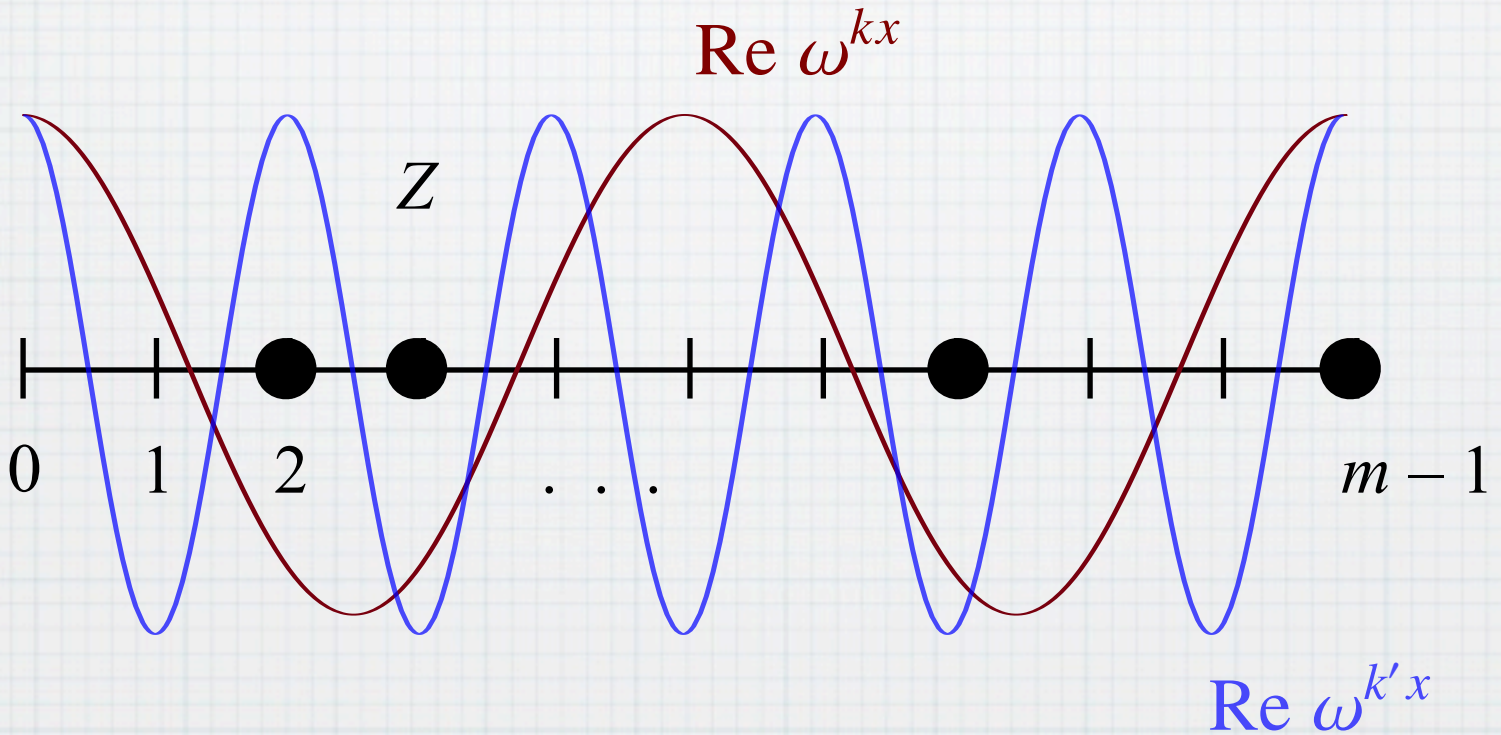
Discrepancy pictorially

Intuition: $\mathrm{disc}_m(Z)$ measures aperiodicity/balancedness



Discrepancy pictorially

Intuition: $\mathrm{disc}_m(Z)$ measures aperiodicity/balancedness



A nonconstructive bound

FACT.

Choose $z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2}\log m\right)} \in \{0, 1, 2, \ldots, m-1\}$ uniformly at random. Then w.h.p.,

$$\operatorname{disc}_m\left(\left\{z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2}\log m\right)}\right\}\right) < \epsilon.$$

A nonconstructive bound

FACT.

Choose $z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2} \log m\right)} \in \{0, 1, 2, \ldots, m-1\}$ uniformly at random. Then w.h.p.,

$$\operatorname{disc}_m\left(\left\{z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2} \log m\right)}\right\}\right) < \epsilon.$$

PROOF: Hoeffding + union bound. □

A nonconstructive bound

tight (simultaneous diophantine approximation)

FACT.

Choose $z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2} \log m\right)} \in \{0, 1, 2, \ldots, m-1\}$ uniformly at random. Then w.h.p.,

$$\operatorname{disc}_m\left(\left\{z_1, z_2, \ldots, z_{O\left(\frac{1}{\epsilon^2} \log m\right)}\right\}\right) < \epsilon.$$

PROOF: Hoeffding + union bound. □

Explicit construction

ITERATION LEMMA (AJTAI ET AL).

Take P, R with $P^2(R+1) \le m$. Fix a set $Z_p \in \{1, 2, \ldots, p-1\}$ for each prime $p \in (P/2, P]$ with $p \nmid m$, such that all Z_p have the same cardinality. Let

$$Z_m = \{(r + s \cdot (p^{-1})_m) \bmod m :$$

$$r = 1, 2, \ldots, R; \quad p \in (P/2, P] \text{ prime with } p \nmid m; \quad s \in Z_p\}.$$

Then

$$\operatorname{disc}(Z_m) = O\left(\frac{1}{\sqrt{R}} + \frac{\ln m}{\ln\ln m} \cdot \frac{\ln P}{P} + \max_p \operatorname{disc}_p(Z_p)\right).$$

Explicit construction

ITERATION LEMMA (AJTAI ET AL).

Take P, R with $P^2(R+1) \le m$. Fix a set $Z_p \in \{1, 2, \ldots, p-1\}$ for each prime $p \in (P/2, P]$ with $p \nmid m$, such that all Z_p have the same cardinality. Let

distinct mod m

$$Z_m = \{(r + s \cdot (p^{-1})_m) \bmod m :$$

$$r = 1, 2, \ldots, R; \quad p \in (P/2, P] \text{ prime with } p \nmid m; \quad s \in Z_p\}.$$

Then

$$\operatorname{disc}(Z_m) = O\left(\frac{1}{\sqrt{R}} + \frac{\ln m}{\ln\ln m} \cdot \frac{\ln P}{P} + \max_p \operatorname{disc}_p(Z_p)\right).$$

Explicit construction

THEOREM (AJTAI ET AL).

For each m, there is an explicit set Z_m with

$$|Z_m| \leq \log^{1+o(1)} m,$$

$$\mathrm{disc}_m(Z_m) = o(1).$$

Explicit construction

THEOREM (AJTAI ET AL).

For each m, there is an explicit set Z_m with

$$|Z_m| \leq \log^{1+o(1)} m,$$

$$\mathrm{disc}_m(Z_m) = o(1).$$

PROOF: recursively apply Iteration Lemma. □

Explicit construction

THEOREM.

For each m and $\epsilon > 0$, there is an explicit set Z_m with

$$|Z_m| = O_\epsilon(\log m),$$

$$\mathrm{disc}_m(Z_m) < \epsilon.$$

Explicit construction

THEOREM.

For each m and $\epsilon > 0$, there is an explicit set Z_m with

$$|Z_m| = O_\epsilon(\log m),$$

$$\mathrm{disc}_m(Z_m) < \epsilon.$$

PROOF: apply Iteration Lemma twice, then use brute force search. □

Step 2:
A random walk on $\mathbb{Z}_m$

A random walk on Z_m

Fix $z_1, z_2, \ldots, z_n \in \{0, 1, 2, \ldots, m-1\}$. *Think* $m = 2^{\Omega(n)}$.

A random walk on Z_m

Fix $z_1, z_2, \ldots, z_n \in \{0, 1, 2, \ldots, m-1\}$. *Think* $m = 2^{\Omega(n)}$.

For $x \in \{0,1\}^n$, analyze distribution of

$$\left(\sum_{i=1}^{n} z_i x_i \right) \bmod m.$$

A random walk on Z_m

Fix $z_1, z_2, \ldots, z_n \in \{0, 1, 2, \ldots, m-1\}$. *Think* $m = 2^{\Omega(n)}$.

For $x \in \{0,1\}^n$, analyze distribution of

$$\left(\sum_{i=1}^{n} z_i x_i \right) \bmod m.$$

n-step random walk on $\mathbf{Z}_m$

A random walk on Z_m

$$\frac{1}{2}\begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} & & & & \overbrace{1}^{z_j \bmod m} & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & & 1 & & & \end{bmatrix}$$

A random walk on Z_m

$$\prod_{i=1}^{n}\left(\frac{1}{2}\begin{bmatrix}1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & 1\end{bmatrix}+\frac{1}{2}\begin{bmatrix} & & & \overbrace{1} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & & 1 & & & \end{bmatrix}\right)$$

A random walk on Z_m

$$\prod_{i=1}^{n}\left(\frac{1}{2}\begin{bmatrix}1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & 1\end{bmatrix}+\frac{1}{2}\begin{bmatrix} & & & \overbrace{1}^{z_j \bmod m} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & & 1 & & & \end{bmatrix}\right)$$

$$= \prod_{i=1}^{n} U \operatorname{diag}\left(1, \frac{1+\omega^{z_i}}{2}, \frac{1+\omega^{2z_i}}{2}, \ldots, \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$

A random walk on Z_m

$$= \prod_{i=1}^{n} U \operatorname{diag}\left(1, \frac{1+\omega^{z_i}}{2}, \frac{1+\omega^{2z_i}}{2}, \ldots, \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$

A random walk on Z_m

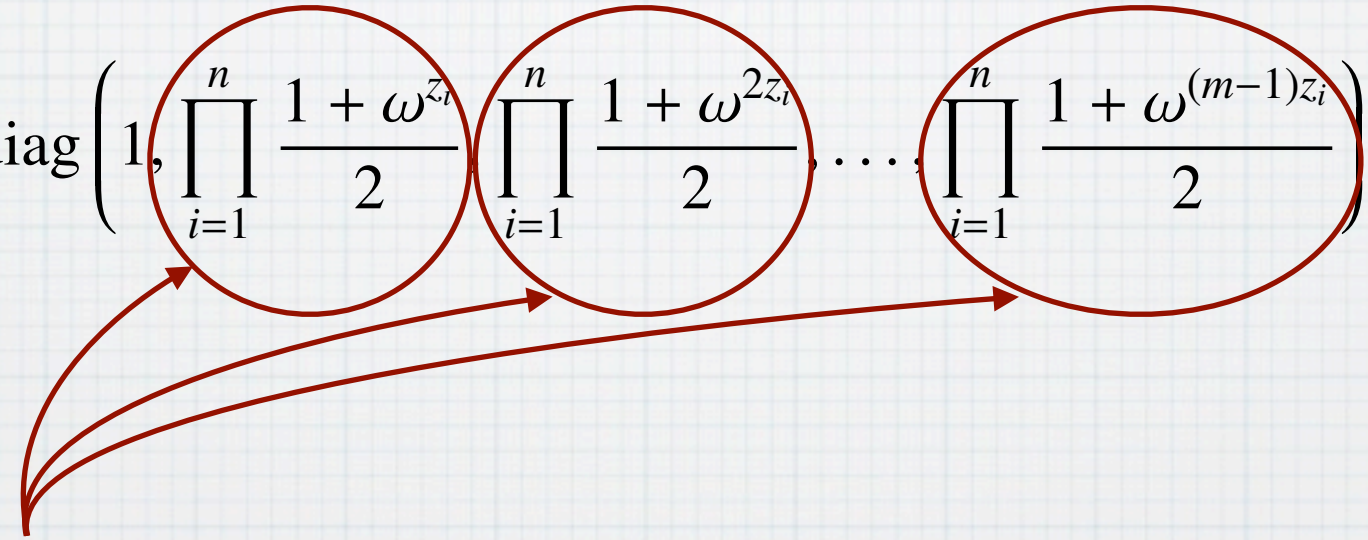
$$= \prod_{i=1}^{n} U \operatorname{diag}\left(1, \frac{1+\omega^{z_i}}{2}, \frac{1+\omega^{2z_i}}{2}, \ldots, \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$

$$= U \operatorname{diag}\left(1, \prod_{i=1}^{n} \frac{1+\omega^{z_i}}{2}, \prod_{i=1}^{n} \frac{1+\omega^{2z_i}}{2}, \ldots, \prod_{i=1}^{n} \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$

A random walk on Z_m

$$= \prod_{i=1}^{n} U \operatorname{diag}\left(1, \frac{1+\omega^{z_i}}{2}, \frac{1+\omega^{2z_i}}{2}, \ldots, \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$

$$= U \operatorname{diag}\left(1, \prod_{i=1}^{n} \frac{1+\omega^{z_i}}{2}, \prod_{i=1}^{n} \frac{1+\omega^{2z_i}}{2}, \ldots, \prod_{i=1}^{n} \frac{1+\omega^{(m-1)z_i}}{2}\right) U^*$$



bound by $\left(\frac{1+\operatorname{disc}_m(\{z_1, z_2, \ldots, z_n\})}{2}\right)^{n/2}$

A random walk on Z_m

This sketches:

LEMMA. For any integers $z_1, z_2, \ldots, z_n, s$,

$$\left| \mathop{\mathbf{P}}_{x \in \{0,1\}^n} \left[\sum_{j=1}^{n} z_j x_j \equiv s \pmod{m} \right] - \frac{1}{m} \right| \leq \left(\frac{1 + \mathrm{disc}_m(Z)}{2} \right)^{n/2}.$$

Step 3:
Fooling distributions

Fooling distributions

Given: $\epsilon = \frac{1}{2018}$

$m = 2^{\epsilon n}$

$z_1, z_2, \ldots, z_n \in \mathbb{Z}_m$ with $\text{disc}_m(\{z_1, z_2, \ldots, z_n\}) < \epsilon$

Fooling distributions

Given: $\epsilon = \frac{1}{2018}$

$m = 2^{\epsilon n}$

$z_1, z_2, \ldots, z_n \in \mathbb{Z}_m$ with $\text{disc}_m(\{z_1, z_2, \ldots, z_n\}) < \epsilon$

Define $L\colon \{0,1\}^n \to \mathbb{Z}_m$ by $L(x) = \sum z_i x_i$

Fooling distributions

Given: $\epsilon = \frac{1}{2018}$

$m = 2^{\epsilon n}$

$z_1, z_2, \ldots, z_n \in \mathbb{Z}_m$ with $\mathrm{disc}_m(\{z_1, z_2, \ldots, z_n\}) < \epsilon$

Define $L\colon \{0,1\}^n \to \mathbb{Z}_m$ by $L(x) = \sum z_i x_i$

By Step 2:

$$\mathop{\mathbf{E}}_{L^{-1}(0)} p \approx \mathop{\mathbf{E}}_{L^{-1}(1)} p \approx \mathop{\mathbf{E}}_{L^{-1}(2)} p \approx \cdots \approx \mathop{\mathbf{E}}_{L^{-1}(m-1)} p$$

for every polynomial p of degree $\leq \varepsilon n$.

Fooling distributions

Given: $\epsilon = \frac{1}{2018}$

$m = 2^{\epsilon n}$

$z_1, z_2, \ldots, z_n \in \mathbb{Z}_m$ with $\mathrm{disc}_m(\{z_1, z_2, \ldots, z_n\}) < \epsilon$

Define $L\colon \{0,1\}^n \to \mathbb{Z}_m$ by $L(x) = \sum z_i x_i$

By Step 2 **and perturbation argument:**

$$\mathop{\mathbf{E}}_{\mu_0} p \approx \mathop{\mathbf{E}}_{\mu_1} p \approx \mathop{\mathbf{E}}_{\mu_2} p \approx \cdots \approx \mathop{\mathbf{E}}_{\mu_{m-1}} p$$

for every polynomial p of degree $\leq \varepsilon n$, **where supp $\mu_i \subseteq L^{-1}(i)$**

Step 4: Univariatization

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
supp $\mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.

$\operatorname{supp} \mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}.$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
$\operatorname{supp} \mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}.$

$$(1 - \delta)Q(x, t) < P(x, t)F(x, t) < (1 + \delta)Q(x, t)$$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
supp $\mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}$.

$$(1-\delta) Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)$$
$$< P\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right) F\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)$$
$$< (1+\delta) Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)$$

$\} x \in \operatorname{supp} \mu_s$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
$\operatorname{supp} \mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}.$

$$(1 - \delta) Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)$$
$$< P\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right) \operatorname{sgn} s$$
$$< (1 + \delta) Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)$$

$x \in \operatorname{supp} \mu_s$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
$\operatorname{supp} \mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}$.

$$(1 - \delta)\, \mathbf{E}\left[Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)\right]$$
$$< \mathbf{E}\left[P\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)\right] \operatorname{sgn} s$$
$$< (1 + \delta)\, \mathbf{E}\left[Q\left(x, 2^{-\varepsilon n} (\textstyle\sum z_i x_i - s)\right)\right]$$

} expectation w.r.t. $x \sim \operatorname{supp} \mu_s$

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
$\operatorname{supp} \mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in \operatorname{supp} \mu_s$:

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}$.

$$(1 - \delta)\, q(s) < p(s) \operatorname{sgn} s < (1 + \delta)\, q(s)$$

for every
$s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.

Univariatization

Define $F : \{0,1\}^n \times \{0, 1, \ldots, n\} \to \{-1,+1\}$ by

$F(x, t) = \operatorname{sgn}(\sum z_i x_i - 2^{\varepsilon n} t)$

halfspace on $\{0,1\}^{2n}$

Fix δ-error approximant P/Q for F
(deg P, deg $Q \le \varepsilon n$; $Q > 0$).

Distributions μ_s for $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.
supp $\mu_s \subseteq \{ x : \sum z_i x_i \equiv s \pmod{2^{\varepsilon n}} \}$

For $x \in$ supp μ_s :

$2^{-\varepsilon n} (\sum z_i x_i - s) \in \{0,1, \ldots, n\}$.

$$(1 - \delta)\, q(s) < p(s) \operatorname{sgn} s < (1 + \delta)\, q(s)$$

for every $s = \pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}$.

p/q approximates sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\varepsilon n - 1}\}$.

Summary

THEOREM.

Any rational approximant of degree $d \leq \varepsilon n$ for F gives a degree-d univariate rational approximant for sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$, with the same error.

Summary

THEOREM.

Any rational approximant of degree $d \le \varepsilon n$ for F gives a degree-d univariate rational approximant for sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$, with the same error.

COROLLARY.

$$E(F, d) \ge 1 - \exp(-\Omega(n)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n,$$
$$R(F, d) \ge 1 - \exp(-\Omega(n / d)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n.$$

Summary

THEOREM.

Any rational approximant of degree $d \le \varepsilon n$ for F gives a degree-d univariate rational approximant for sgn on $\{\pm 1, \pm 2, \ldots, \pm 2^{\Omega(n)}\}$, with the same error.

COROLLARY.

$$E(F, d) \ge 1 - \exp(-\Omega(n)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n,$$
$$R(F, d) \ge 1 - \exp(-\Omega(n / d)) \quad \text{for } d = 0, 1, \ldots, \varepsilon n.$$

COROLLARY. $E(F \wedge F, \Omega(n)) = 1,$
$R(F \wedge F, \Omega(n)) = 1.$

Questions?