Sublinear Algorithms

LECTURE 5

Last time

- Limitations of sublinear-time algorithms
- Yao's Minimax Principle
    - Examples: testing 0^* and sortedness

Today

- Limitations of sublinear-time algorithms
- Yao's Minimax Principle
- Communication complexity

HW1 resubmission, HW3 out, project guidelines

Recall: Yao's Minimax Principle

Statement 1

For any **probabilistic** algorithm A of complexity q there exists an input x s.t.

$$\Pr_{coin\ tosses\ of\ A}[\text{A(x) is wrong}] > 1/3.$$

Statement 2

There is a distribution *D* on the inputs,

s.t. for every **deterministic** algorithm of complexity q,

$$\Pr_{x \leftarrow D}[\text{A(x) is wrong}] > 1/3.$$

- Need for lower bounds

Yao's Minimax Principle (easy direction): Statement 2 $\Rightarrow$ Statement 1.

NOTE: Also applies to restricted algorithms

- 1-sided error tests
- nonadaptive tests

Yao's Minimax Principle as a game

Players: Evil algorithms designer Al and poor lower bound prover Lola.

Game1

Move 1. Al selects a q-query **randomized** algorithm A for the problem.

Move 2. Lola selects an input on which A errs with largest probability.

Game2

Move 1. Lola selects a distribution on inputs.

Move 2. Al selects a q-query **deterministic** algorithm with as large probability of success on Lola's distribution as possible.

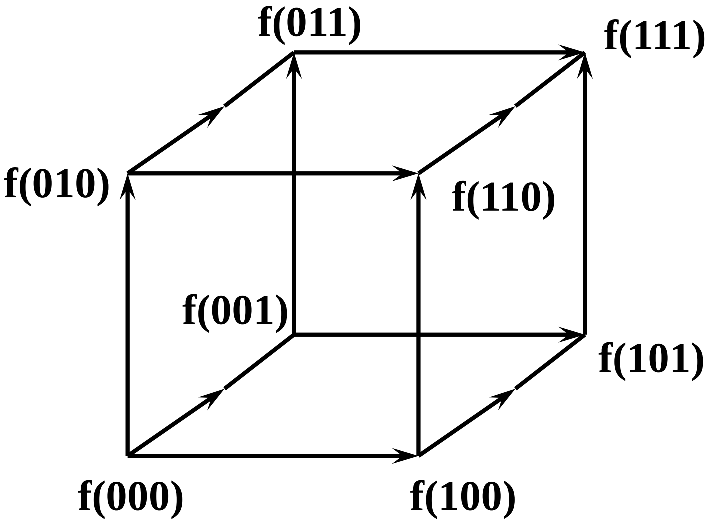
Testing Monotonicity of functions on Hypercube

Non-adaptive 1-sided error Lower Bound

Boolean Functions $f : \{0,1\}^n \to \{0,1\}$

Graph representation:

n-dimensional hypercube



- vertices: bit strings of length n
- edges: (x, y) is an edge if y can be obtained from x by increasing one bit from 0 to 1

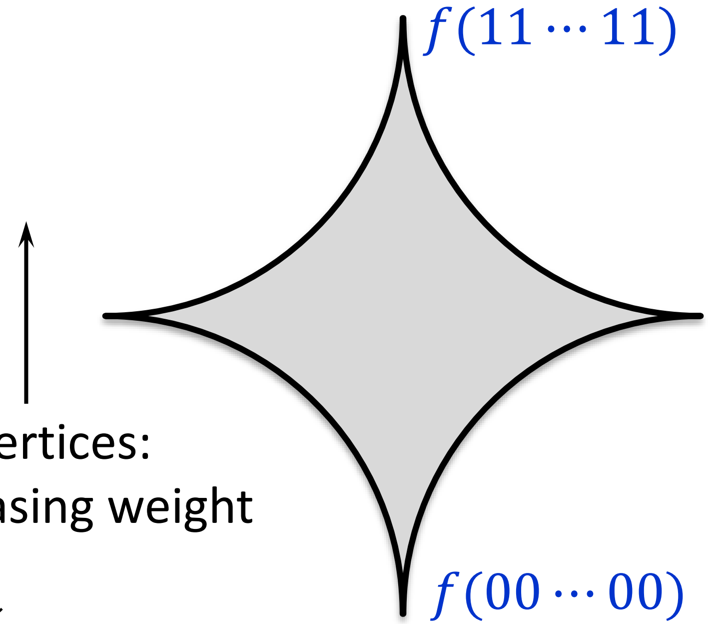
x	001001
y	011001

- each vertex x is labeled with $f(x)$

Boolean Functions $f: \{0,1\}^n \to \{0,1\}$

Graph representation:

n-dimensional hypercube



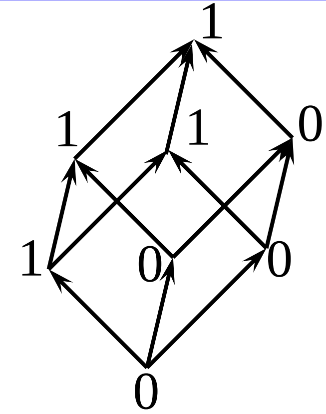
- 2^n vertices: bit strings of length n
- $2^{n-1}n$ edges: (x, y) is an edge if y can be obtained from x by increasing one bit from 0 to 1

x	001001
y	011001

- each vertex x is labeled with $f(x)$

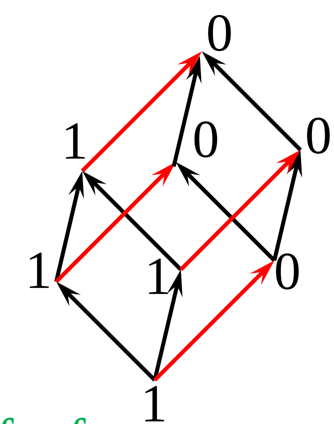
Monotonicity of Functions

[Goldreich Goldwasser Lehman Ron Samorodnitsky,
Dodis Goldreich Lehman Raskhodnikova Ron Samorodnitsky
Fischer Lehman Newman Raskhodnikova Rubinfeld Samorodnitsky]



- A function $f : \{0,1\}^n \to \{0,1\}$ is monotone if increasing a bit of x does not decrease $f(x)$.

- Is f monotone or ε-far from monotone (f has to change on many points to become monontone)?
  - Edge $x \to y$ is violated by f if $f(x) > f(y)$.



Time:

- $O(n/\varepsilon)$, logarithmic in the size of the input, 2^n
- $\Omega(\sqrt{n}/\varepsilon)$ for 1-sided error, nonadaptive tests
- Advanced techniques: $\Theta(\sqrt{n}/\varepsilon^2)$ for nonadaptive tests, $\Omega(\sqrt[3]{n})$

  [Khot Minzer Safra 15, Chen De Servidio Tang 15, Chen Waingarten Xie 17]

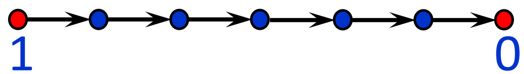
Hypercube 1-sided Error Lower Bound

Lemma [Fischer Lehman Newman Raskhodnikova Rubinfeld Samorodnitsky]

Every 1-sided error nonadaptive test for monotonicity of functions $f : \{0,1\}^n \to \{0,1\}$ requires $\Omega(\sqrt{n})$ queries.

- 1-sided error test must accept if no violated pair is uncovered.

Violated pair:

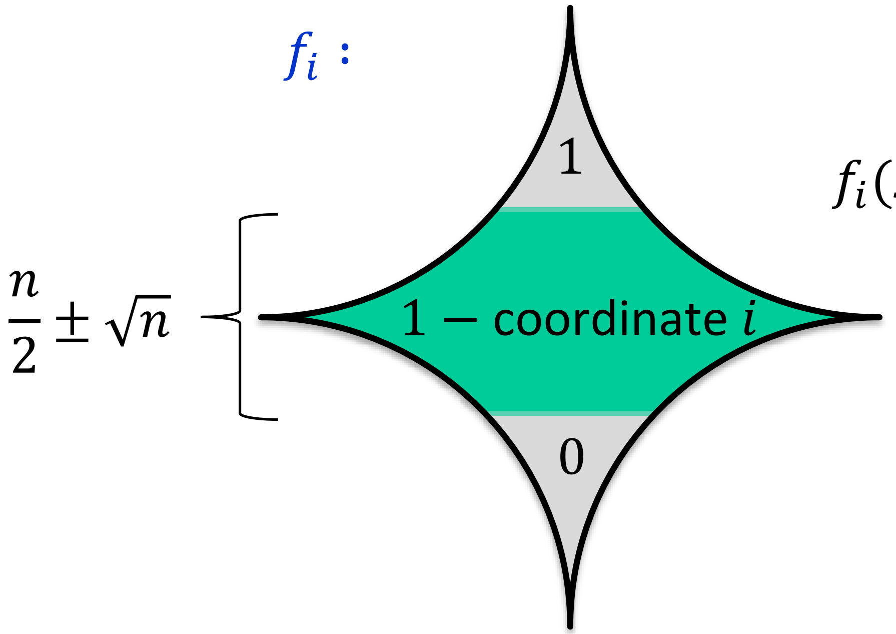


  - A distribution on far from monotone functions suffices.

Hypercube 1-sided Error Lower Bound

- Hard distribution: pick coordinate i at random and output f_i.

f_i :

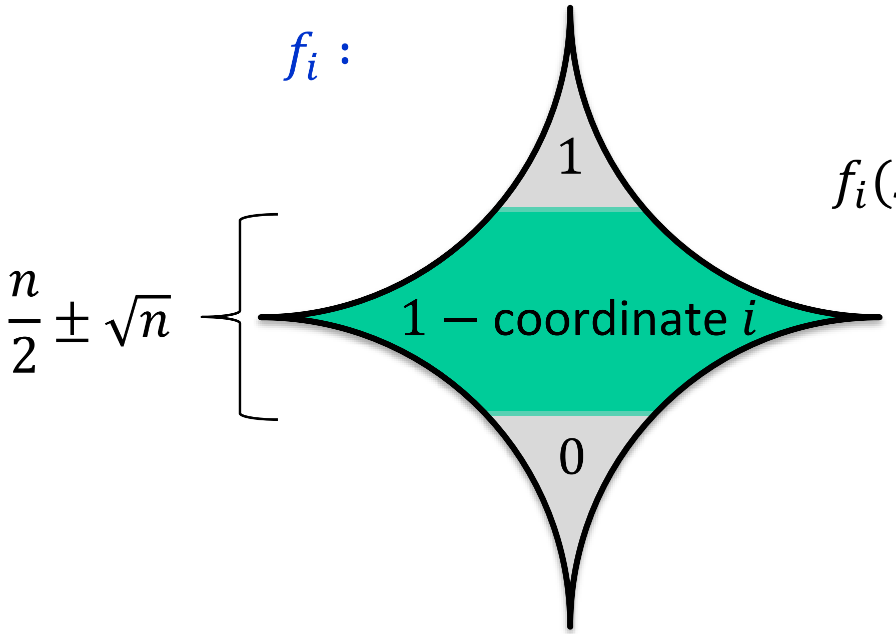


$$f_i(x) = \begin{cases} 1 & \text{if } |x| > \frac{n}{2} + \sqrt{n} \\ 1 - x_i & \text{if } |x| = \frac{n}{2} \pm \sqrt{n} \\ 0 & \text{if } |x| < \frac{n}{2} - \sqrt{n} \end{cases}$$

Hypercube 1-sided Error Lower Bound

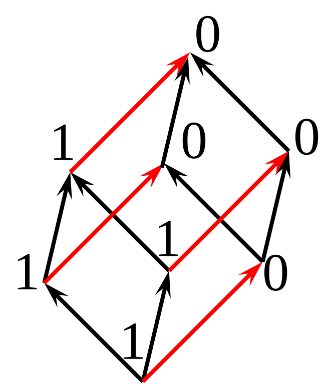
- Hard distribution: pick coordinate i at random and output f_i.

f_i :



$$f_i(x) = \begin{cases} 1 & \text{if } |x| > \frac{n}{2} + \sqrt{n} \\ 1 - x_i & \text{if } |x| = \frac{n}{2} \pm \sqrt{n} \\ 0 & \text{if } |x| < \frac{n}{2} - \sqrt{n} \end{cases}$$

- A ``truncation'' of an antidictator



antidictator

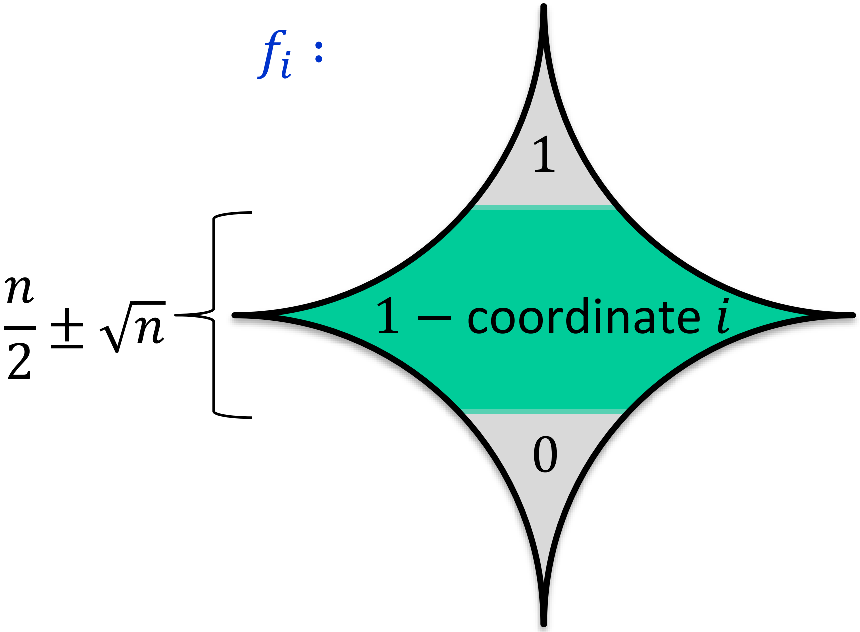
The Fraction of Nodes in Middle Layers

Hoeffding Bound

Let $Y_1, \ldots, Y_s$ be independently distributed random variables in [0,1].

Let $Y = \frac{1}{s} \cdot \sum_{i=1}^{s} Y_i$ (called *sample mean*). Then $\Pr[|Y - E[Y]| \geq \varepsilon] \leq 2e^{-2s\varepsilon^2}$.

f_i :



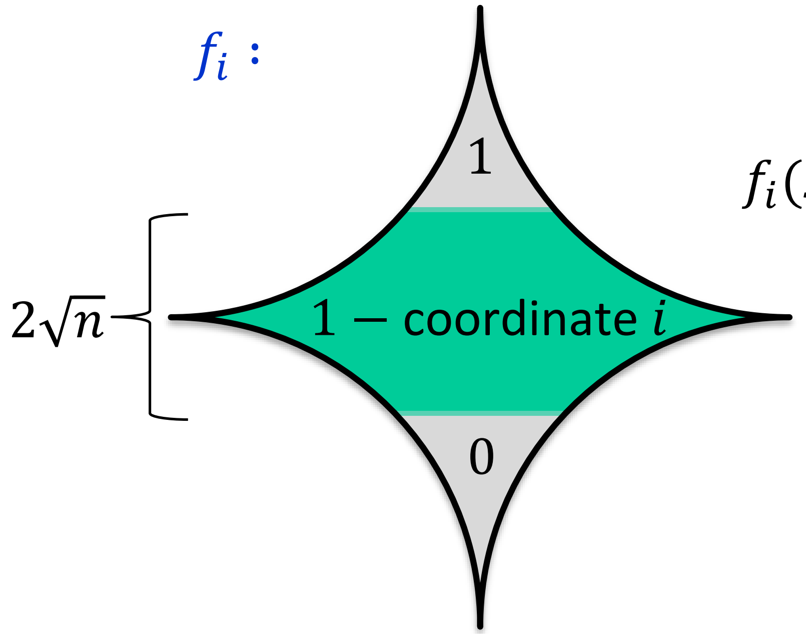
E[Y]=

$\varepsilon =$

Hard Functions are Far

- Hard distribution: pick coordinate i at random and output f_i.

f_i :



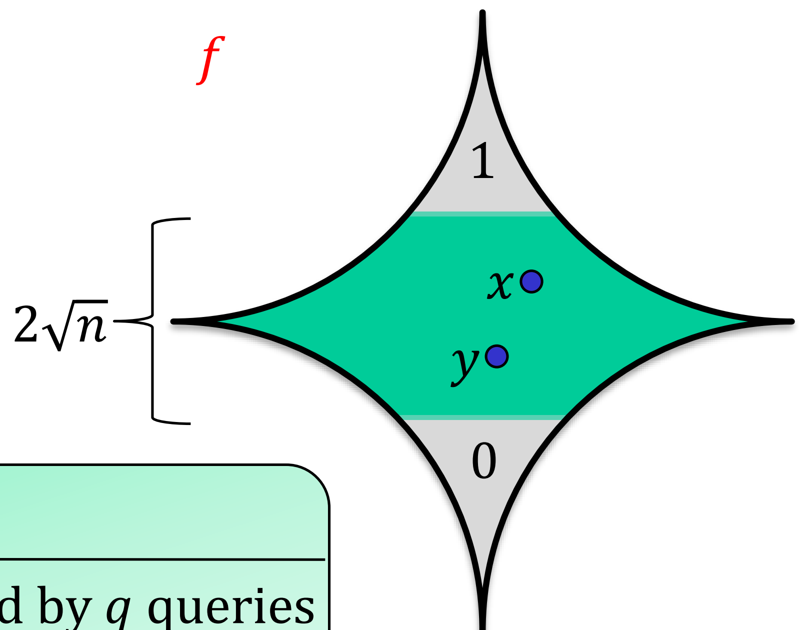
$$f_i(x) = \begin{cases} 1 & \text{if } |x| > \frac{n}{2} + \sqrt{n} \\ 1 - x_i & \text{if } |x| = \frac{n}{2} \pm \sqrt{n} \\ 0 & \text{if } |x| < \frac{n}{2} - \sqrt{n} \end{cases}$$

Analysis

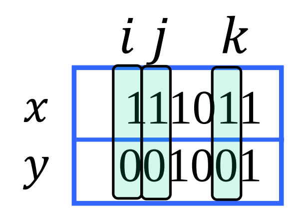
- The middle contains a constant fraction of vertices.
- Edges from $(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n)$ to $(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n)$ are violated if both endpoints are in the middle.
- All n functions are ε-far from monotone for some constant ε.

Hypercube 1-sided Error Lower Bound

- How many functions does a set of q queries expose?



● queries



Pair (x, y)
can expose only
functions f_i, f_j and f_k

Naive Analysis

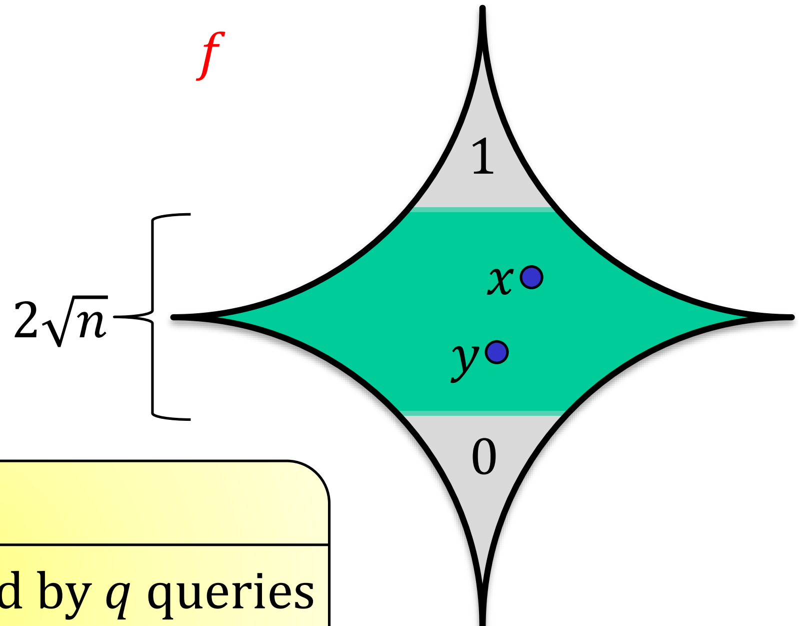
functions exposed by q queries
$\leq q^2 \cdot 2\sqrt{n}$

functions that a query pair (x, y) exposes
$\leq$ # coordinates on which x and y differ
$\leq 2\sqrt{n}$

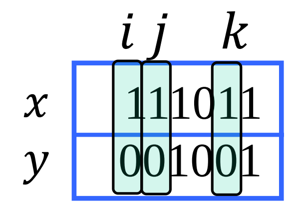
Only pairs of queries in the Green Band can be violated $\Rightarrow$ disagreements $\leq 2\sqrt{n}$

Hypercube 1-sided Error Lower Bound

- How many functions does a set of q queries expose?



● queries



Pair (x, y)
can expose only
functions f_i, f_j and f_k

Claim

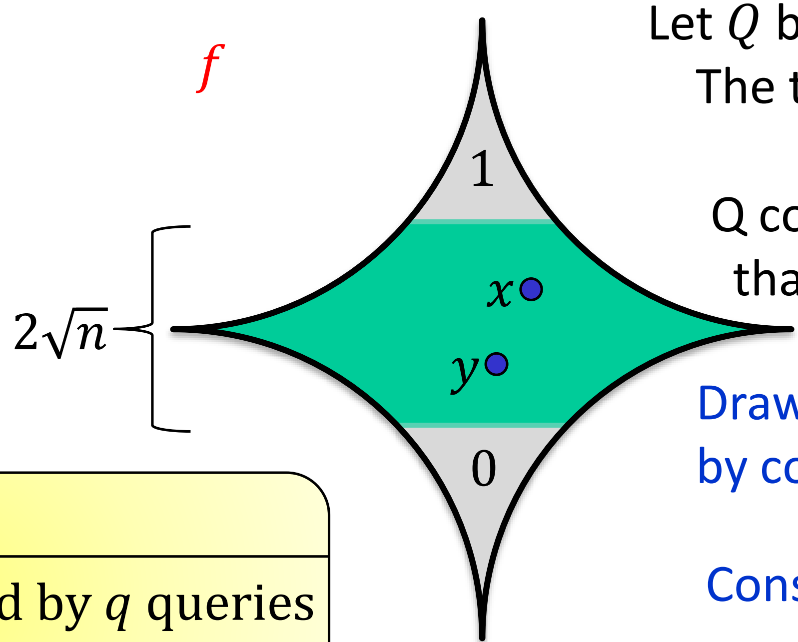
\# functions exposed by q queries
$\leq (q-1) \cdot 2\sqrt{n}$

\# functions that a query pair (x, y) exposes
$\leq$ # coordinates on which x and y differ
$\leq 2\sqrt{n}$

Only pairs of queries in the Green Band can be violated $\Rightarrow$ disagreements $\leq 2\sqrt{n}$

Hypercube 1-sided Error Lower Bound

- How many functions does a set of q queries expose?



Let Q be the set of queries made.
The tester catches a violation
$\Updownarrow$
Q contains comparable x, y
that differ in coordinate i

Draw an undirected graph (Q, E)
by connected comparable queries

Consider its spanning forest.

Claim

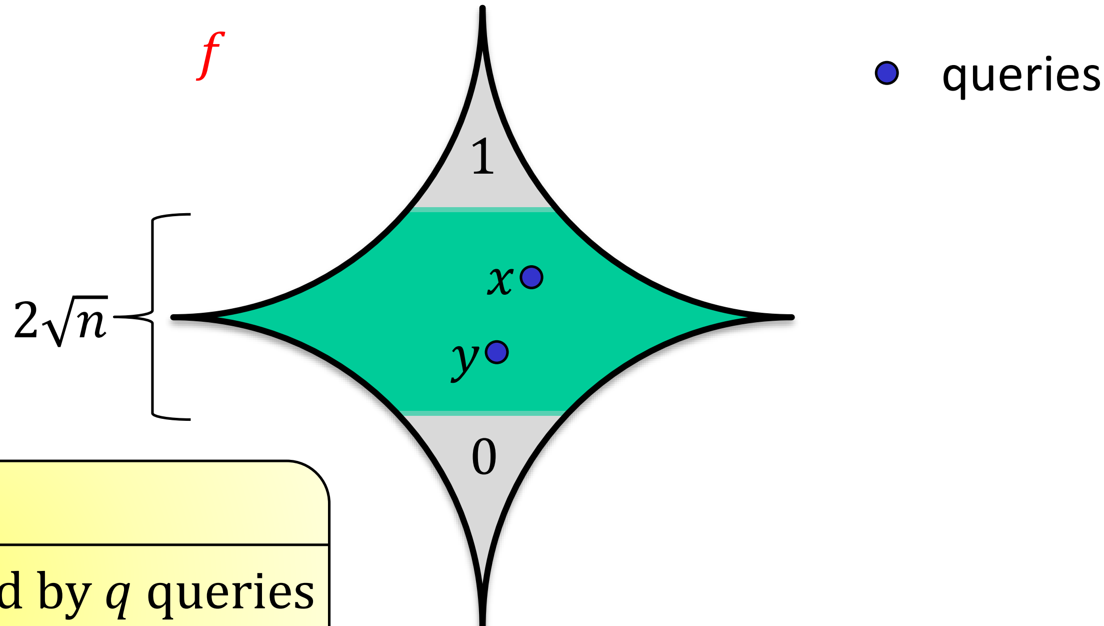
functions exposed by q queries
$\leq (q-1) \cdot 2\sqrt{n}$

sufficient to consider adjacent vertices in a minimum spanning forest on the query set

x, y exist
$\Updownarrow$
there are adjacent vertices on the path from x to y that differ in coordinate i

Hypercube 1-sided Error Lower Bound

- How many functions does a set of q queries expose?



Claim

\# functions exposed by q queries $\leq (q-1) \cdot 2\sqrt{n}$

$\Downarrow$

Claim

Every deterministic test that makes a set Q of q queries (in the middle) succeeds with probability $O\left(\frac{q}{\sqrt{n}}\right)$ on our distribution.

Communication Complexity

A Method for Proving Lower Bounds

[Blais Brody Matulef 11]

Use known lower bounds
for other models of computation

Partially based on slides by Eric Blais

(Randomized) Communication Complexity

Shared random string

Alice | 11010001011101011101010101110... | Bob

0100 →

← 11

001 →

…

← 0011

Input: x

Input: y

Compute $C(x, y)$

Goal: minimize the number of bits exchanged.

- Communication complexity of a protocol is the maximum number of bits exchanged by the protocol.
- Communication complexity of a function C, denoted $R(C)$, is the communication complexity of the best protocol for computing C.

Example: Set Disjointness $DISJ_k$

Alice

110100010111010111010101010110...

Bob

Input: $S \subseteq [n]$, $|S| = k$.

Input: $T \subseteq [n]$, $|T| = k$

Compute $DISJ_k(S,T)$

$$= \begin{cases} \textbf{accept} & \text{if } S \cap T = \emptyset \\ \textbf{reject} & \text{otherwise} \end{cases}$$

Theorem [Kalyanasundaram Schmitger 92, Razborov 92]

$R(\mathrm{DISJ}_k) \geq \Omega(k)$ for all $k \leq \frac{n}{2}$.

A lower bound using CC method

Testing if a Boolean function is a k-parity

Linear Functions Over Finite Field $\mathbb{F}_2$

A Boolean function $f: \{0,1\}^n \to \{0,1\}$ is *linear* (also called *parity*) if

$$f(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n \text{ for some } a_1, \dots, a_n \in \{0,1\}$$

no free term

- Work in finite field $\mathbb{F}_2$
  - Other accepted notation for $\mathbb{F}_2$: GF_2 and $\mathbb{Z}_2$
  - Addition and multiplication is mod 2
  - $\boldsymbol{x}=(x_1, \dots, x_n)$, $\boldsymbol{y}=(y_1, \dots, y_n)$, that is, $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$

    $\boldsymbol{x} + \boldsymbol{y}=(x_1 + y_1, \dots, x_n + y_n)$

example

$$\begin{array}{r} 001001 \\ +\ 011001 \\ \hline 010000 \end{array}$$

Linear Functions Over Finite Field $\mathbb{F}_2$

A Boolean function $f: \{0,1\}^n \to \{0,1\}$ is *linear* (also called *parity*) if

$$f(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n \text{ for some } a_1, \dots, a_n \in \{0,1\}$$

$$\Updownarrow$$

$$f(x_1, \dots, x_n) = \sum_{i \in S} x_i \textit{ for some } S \subseteq [n].$$

$[n]$ is a shorthand for $\{1, \dots n\}$

Notation: $\chi_S(x) = \sum_{i \in S} x_i$.

Testing if a Boolean function is Linear

Input: Boolean function $f: \{0,1\}^n \rightarrow \{0,1\}$

Question:

Is the function linear or ε-far from linear

($\geq \varepsilon 2^n$ values need to be changed to make it linear)?

Later in the course:

Famous BLR (Blum Lubi Rubinfeld 90) test runs in $O\left(\frac{1}{\varepsilon}\right)$ time

k-Parity Functions

k-Parity Functions

A function $f : \{0,1\}^n \to \{0,1\}$ is a k-parity if

$$f(x) = \chi_S(x) = \sum_{i \in S} x_i$$

for some set $S \subseteq [n]$ of size $|S| = k$.

Testing if a Boolean Function is a k-Parity

Input: Boolean function $f: \{0,1\}^n \to \{0,1\}$ and an integer k

Question: Is the function a k-parity or ε-far from a k-parity ($\geq \varepsilon 2^n$ values need to be changed to make it a k-parity)?

Time:

$\mathrm{O}(k \log k)$ [Chakraborty Garcia−Soriano Matsliah]

$\Omega(\min(k, n-k))$ [Blais Brody Matulef 11]

- Today: $\Omega(k)$ for $k \leq n/2$

Today's bound implies $\Omega(\min(k, n-k))$

Important Fact About Linear Functions

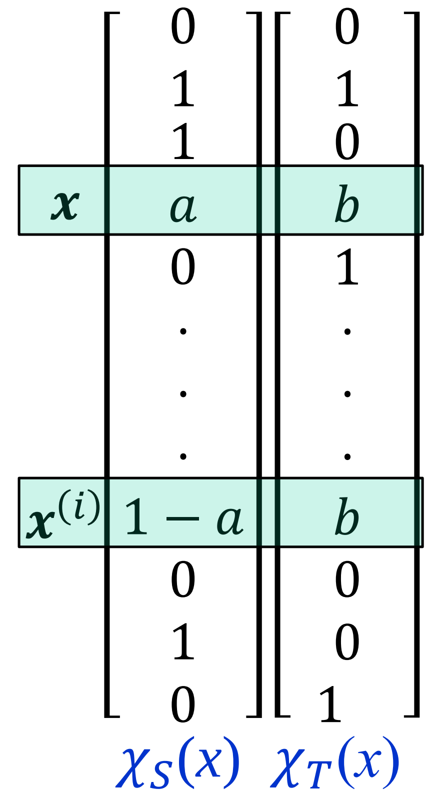
Fact. Two different linear functions disagree on half of the values.

- Consider functions χ_S and χ_T where $S \neq T$.
  - Let i be an element on which S and T differ (w.l.o.g. $i \in S \setminus T$)
  - Pair up all n-bit strings: $(\boldsymbol{x}, \boldsymbol{x}^{(i)})$ where $\boldsymbol{x}^{(i)}$ is $\boldsymbol{x}$ with the i^{th} bit flipped.
  - For each such pair, $\chi_S(\boldsymbol{x}) \neq \chi_S(\boldsymbol{x}^{(i)})$

    but $\chi_T(\boldsymbol{x}) = \chi_T(\boldsymbol{x}^{(i)})$

    So, χ_S and χ_T differ on exactly one of $\boldsymbol{x}, \boldsymbol{x}^{(i)}$.
  - Since all $\boldsymbol{x}$'s are paired up,

    χ_S and χ_T differ on half of the values.



Corollary. A k'−parity function, where $k' \neq k$, is ½-far from any k-parity.

Reduction from $DISJ_{\boldsymbol{k/2}}$ *to Testing k-Parity*

- Let T be the best tester for the k-parity property for $\varepsilon = 1/2$
  - query complexity of T is q (testing k−parity).
- We will construct a communication protocol for $DISJ_{\boldsymbol{k/2}}$ that runs T and has communication complexity $2 \cdot q$(testing k−parity).

holds for CC of every protocol for $DISJ_{\boldsymbol{k}}$

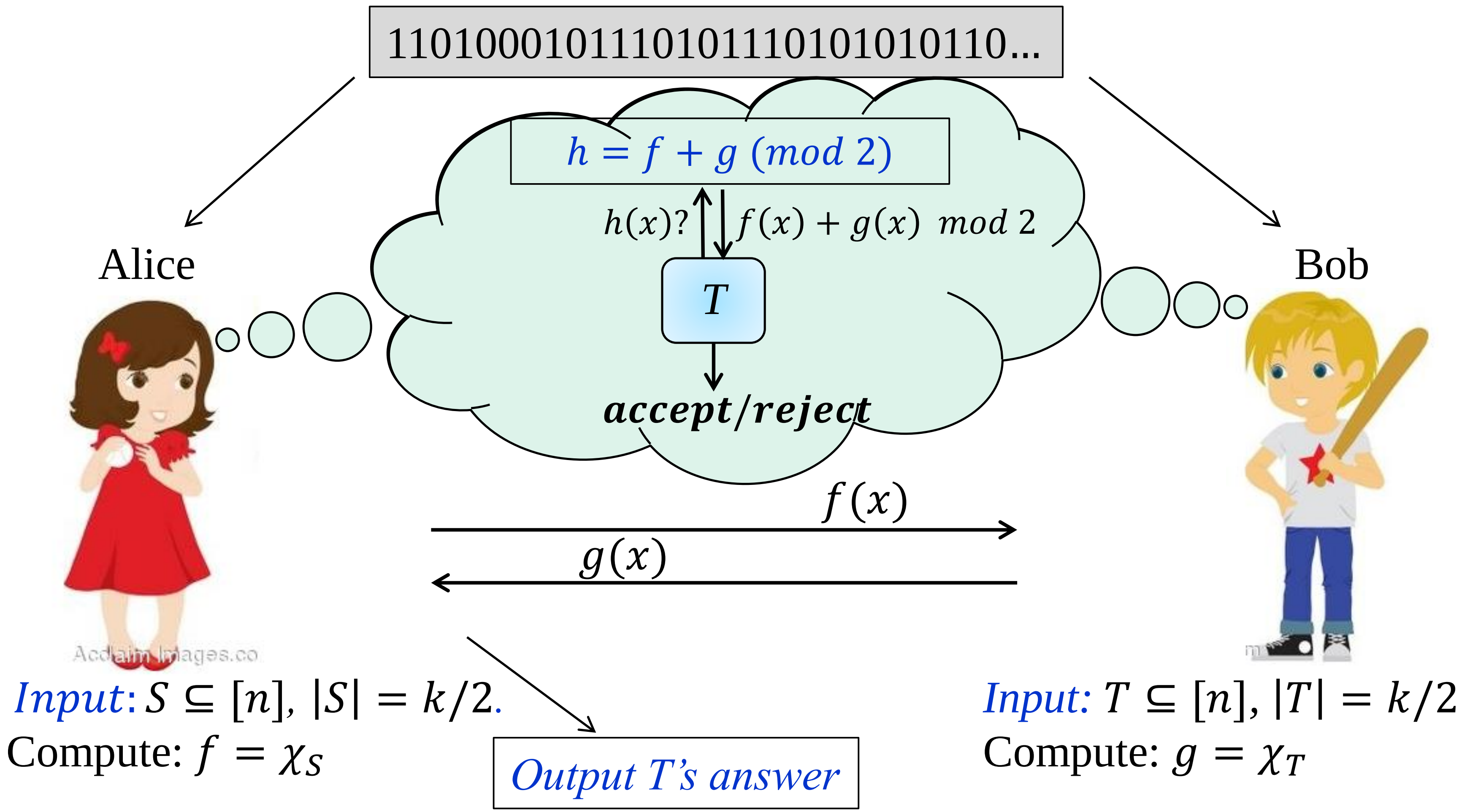
[Kalyanasundaram Schnitger 92]

- Then $2 \cdot q$(testing k−parity) $\geq R\left(\mathrm{DISJ}_{k/2}\right) \geq \Omega(k/2)$ for $k \leq n/2$

$$\Downarrow$$

q(testing k-parity) $\geq \Omega(k)$ for $k \leq n/2$

Reduction from $DISJ_{k/2}$ to Testing k-Parity



- T receives its random bits from the shared random string.

Analysis of the Reduction

Queries: Alice and Bob exchange 2 bits for every bit queried by T

Correctness:

- $h = f + g \ (mod\ 2) = \chi_S + \chi_T \ (mod\ 2) = \chi_{S\Delta T}$
- $|S\Delta T| = |S| + |T| - 2|S \cap T|$

- $|S\Delta T| = \begin{cases} k & \text{if } S\cap T = \emptyset \\ \leq k-2 & \text{if } S\cap T \neq \emptyset \end{cases}$

$$h \text{ is } \begin{cases} k\text{−parity} & \text{if } S\cap T = \emptyset \\ k'\text{−parity where } k' \neq k & \text{if } S\cap T \neq \emptyset \end{cases}$$

1/2-far from every k-parity

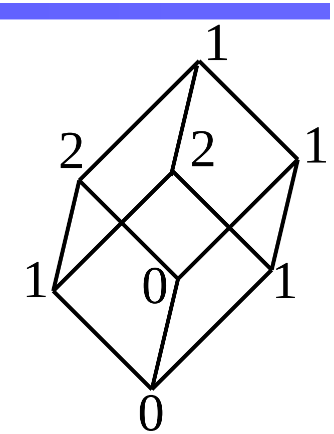
Summary: q(testing k-parity) $\geq \Omega(k)$ for $k \leq n/2$

Testing Lipschitz Property on Hypercube

Lower Bound

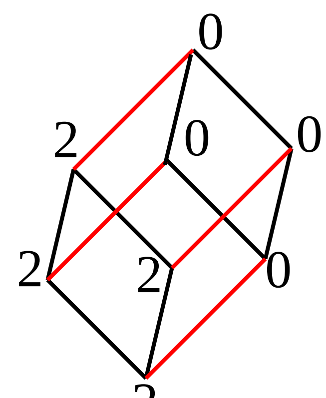
Lipschitz Property of Functions f: $\{0,1\}^n \to \boldsymbol{R}$

- A function $f : \{0,1\}^n \to \mathrm{R}$ is Lipschitz
  if changing a bit of x changes $f(x)$ by at most 1.



Lipschitz

- Is f Lipschitz or ε-far from Lipschitz
  (f has to change on many points to become Lipschitz)?
  - Edge $x - y$ is violated by f if $|f(x) - f(y)| > 1$.



$\frac{1}{2}$-far from Lipschitz

Time:

- $O(n/\varepsilon)$, logarithmic in the size of the input, 2^n
  [Chakrabarty Seshadhri]
- $\Omega(n)$ [Jha Raskhodnikova]

Testing Lipschitz Property

Theorem

Testing Lipschitz property of functions f: $\{0,1\}^n \rightarrow \{0,1,2\}$ requires $\Omega(n)$ queries.

Prove it.

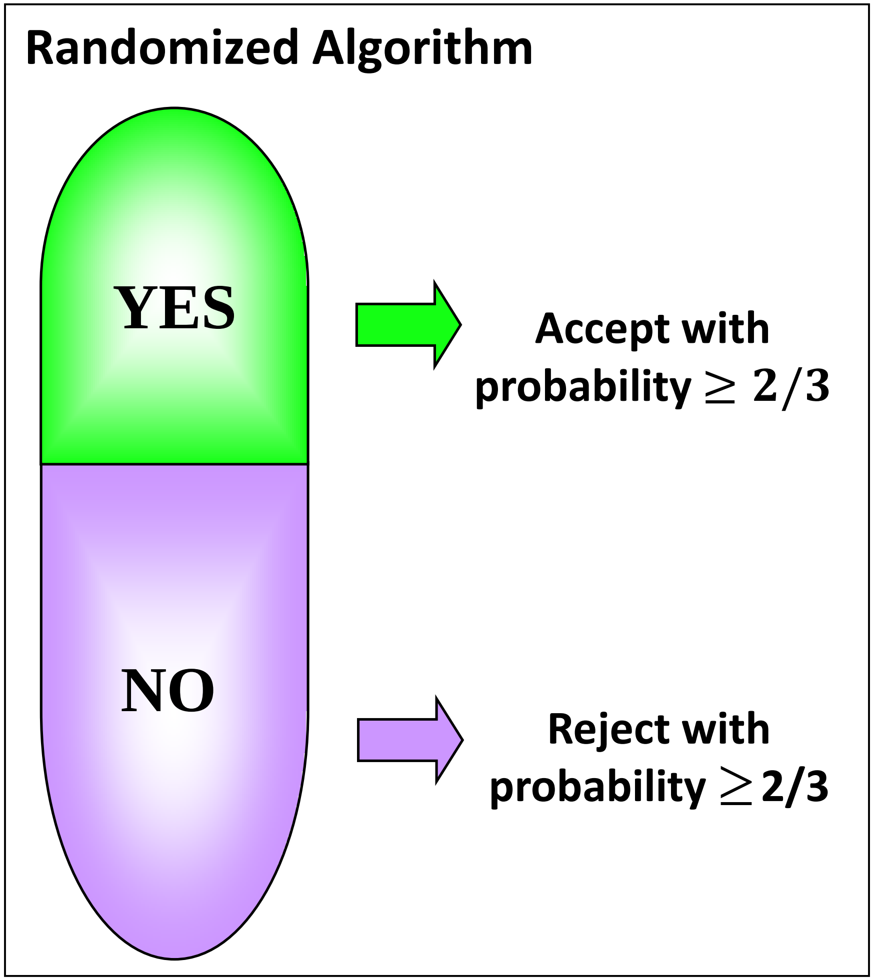
Summary of Lower Bound Methods

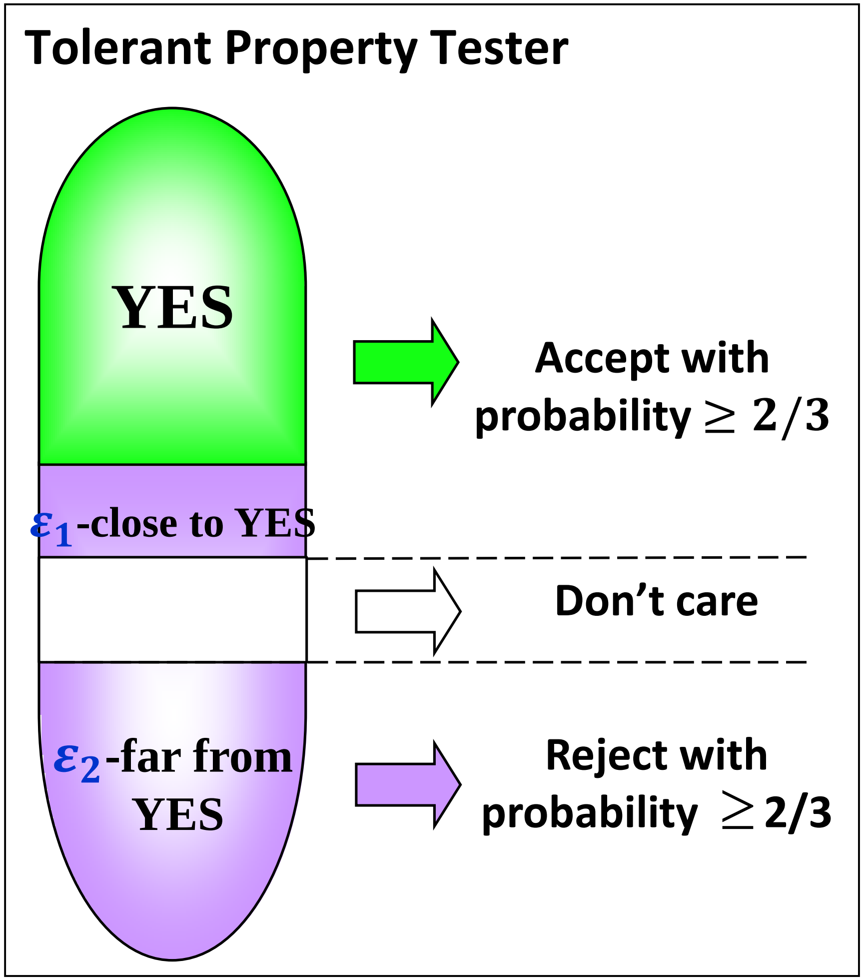
- Yao's Principle
  - testing membership in 1*, sortedness of a list and monotonicity of Boolean functions

- Reductions from communication complexity problems
  - testing if a Boolean function is a k-parity

Other Models of Sublinear Computation

Tolerant Property Tester [Rubinfeld Parnas Ron]



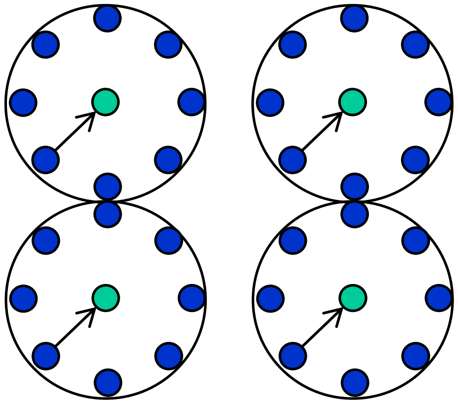


Sublinear-Time "Restoration" Models

Local Decoding

Input: A slightly corrupted codeword

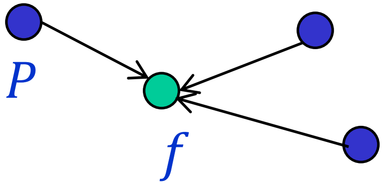
Requirement: Recover individual bits of the closest codeword with a constant number of queries per recovered bit.



Program Checking

Input: A program P computing f correctly on most inputs.

Requirement: Self-correct program P: for a given input x, compute $f(x)$ by making a few calls to *P*.



Local Reconstruction

Input: Function f nearly satisfying some property P

Requirement: Reconstruct function f to ensure that the reconstructed function g satisfies P, changing f only when necessary. For each input x, compute $g(x)$ with a few queries to f.