Sublinear Algorithms

LECTURE 18

Last time

- Finish approximating the average degree
- Testing linearity of Boolean functions

Today

- Finish testing linearity of Boolean functions

  [Blum Luby Rubinfeld]
- Tolerant testing and distance estimation

HW 4 is due Thursday

Next week: *Fourier-Monte Carlo Descent Trees, which combine harmonic analysis with stochastic branch pruning to estimate edge-connectivity in near-quantum time.*

Testing If a Boolean Function Is Linear

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

Question:

Is the function linear or ε-far from linear

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

Today: can answer in $O\left(\frac{1}{\varepsilon}\right)$ time

Linearity Test [Blum Luby Rubinfeld 90]

BLR Test (ε, query access to f)

1. Pick $\boldsymbol{x}$ and $\boldsymbol{y}$ independently and uniformly at random from $\{0,1\}^n$.
2. Set $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$ and query f on $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$. **Accept** iff $f(\boldsymbol{z}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

Analysis

If f is linear, BLR always accepts.

Correctness Theorem [Bellare Coppersmith Hastad Kiwi Sudan 95]

If f is ε-far from linear then $> \varepsilon$ fraction of pairs $\boldsymbol{x}$ and $\boldsymbol{y}$ fail BLR test.

- Then, by Witness Lemma (Lecture 1), $2/\varepsilon$ iterations suffice.

Analysis Technique: Fourier Expansion

Representing Functions as Vectors

Stack the 2^n values of $f(\boldsymbol{x})$ and treat it as a vector in $\{0,1\}^{2^n}$.

$$f = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} f(0000) \\ f(0001) \\ f(0010) \\ f(0011) \\ f(0100) \\ \cdot \\ \cdot \\ \cdot \\ f(1101) \\ f(1110) \\ f(1111) \end{bmatrix}$$

Linear functions

There are 2^n linear functions: one for each subset $S \subseteq [n]$.

$$\chi_{\emptyset} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \chi_{\{1\}} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \cdots \cdots, \qquad \chi_{[n]} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Parity on the positions indexed by set S is $\chi_S(x_1, \dots, x_n) = \sum_{i \in S} x_i$

Great Notational Switch

Idea: Change notation, so that we work over reals instead of a finite field.

- Vectors in $\{0,1\}^{2^n}$ $\longrightarrow$ Vectors in $\mathbb{R}^{2^n}$.
- 0/False $\longrightarrow$ 1 1/True $\longrightarrow$ -1.
- Addition (mod 2) $\longrightarrow$ Multiplication in $\mathbb{R}$.
- Boolean function: $f : \{-1, 1\}^n \to \{-1,1\}$.
- Linear function χ_S: $\{-1, 1\}^n \to \{-1,1\}$ is given by $\chi_S(\boldsymbol{x}) = \prod_{i \in S} x_i$.

Benefits of New Notation

Inner product of functions $f, g : \{-1,1\}^n \to \{-1,1\}$

$$\langle f, g \rangle = \frac{1}{2^n}\text{(dot product of } f \text{ and } g \text{ as vectors)}$$

$$= \underset{\boldsymbol{x}\in\{-1,1\}^n}{\text{avg}}[f(\boldsymbol{x})g(\boldsymbol{x})] = \underset{\boldsymbol{x}\in\{-1,1\}^n}{\mathbb{E}}[f(\boldsymbol{x})g(\boldsymbol{x})].$$

$\langle f, g \rangle = 1 - 2 \cdot$ (fraction of *disagreements* between f and g)

Claim. The functions $(\chi_S)_{S\subseteq[n]}$ form an orthonormal basis for $\mathbb{R}^{2^n}$.

Fourier Expansion Theorem

Idea: Work in the basis $(\chi_S)_{S\subseteq[n]}$, so it is easy to see how close a specific function f is to each of the linear functions.

Fourier Expansion Theorem

Every function $f : \{-1,1\}^n \to \mathbb{R}$ is uniquely expressible as a linear combination (over $\mathbb{R}$) of the 2^n linear functions:

$$f = \sum_{S\subseteq[n]} \hat{f}(S)\chi_S,$$

where $\hat{f}(S) = \langle f, \chi_S\rangle$ is the Fourier Coefficient of f on set S.

Parseval Equality

Parseval Equality

Let $f : \{-1,1\}^n \to \mathbb{R}$. Then

$$\langle f, f \rangle = \sum_{S \subseteq [n]} \hat{f}(S)^2$$

Proof:

By Fourier Expansion Theorem

$$\langle f, f \rangle = \left\langle \sum_{S \subseteq [n]} \hat{f}(S)\chi_S , \sum_{T \subseteq [n]} \hat{f}(T)\chi_T \right\rangle$$

By linearity of inner product

$$= \sum_S \sum_T \hat{f}(S)\, \hat{f}(T) \langle \chi_S, \chi_T \rangle$$

By orthonormality of χ_S's

$$= \sum_S \hat{f}(S)^2$$

Parseval Equality

Parseval Equality for Boolean Functions

Let $f: \{-1,1\}^n \to \{-1,1\}$. Then

$$\langle f, f \rangle = \sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$$

Proof:

By definition of inner product

$$\langle f, f \rangle = \mathop{\mathbb{E}}_{\boldsymbol{x} \in \{-1,1\}^n}[f(\boldsymbol{x})^2]$$

Since f is Boolean

$$= 1$$

BLR Test in {-1,1} Notation

BLR Test (f, ε)

1. Pick $\boldsymbol{x}$ and $\boldsymbol{y}$ independently and uniformly at random from $\{-1,1\}^n$.
2. Set $\boldsymbol{z} = \boldsymbol{x} \circ \boldsymbol{y}$ and query f on $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$. **Accept** iff $f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z}) = 1$.

Vector product notation: $\boldsymbol{x} \circ \boldsymbol{y} = (x_1 y_1, x_2 y_2, \ldots, x_n y_n)$

Sum-Of-Cubes Lemma. $\Pr_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[\text{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{f}(S)^3$

Proof: Indicator variable $\mathbb{1}_{BLR} = \begin{cases} 1 & \text{if BLR accepts} \\ 0 & \text{otherwise} \end{cases}$

$$\mathbb{1}_{BLR} = \frac{1}{2} + \frac{1}{2} f(\boldsymbol{x}) f(\boldsymbol{y}) f(\boldsymbol{z}).$$

$$\Pr_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^n}[\text{BLR}(f)\text{accepts}] = \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[\mathbb{1}_{BLR}] = \frac{1}{2} + \frac{1}{2} \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[f(\boldsymbol{x}) f(\boldsymbol{y}) f(\boldsymbol{z})]$$

By linearity of expectation

Proof of Sum-Of-Cubes Lemma

So far: $\Pr_{x,y \in \{-1,1\}^n}[\text{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2} \mathop{\mathrm{E}}_{x,y \in \{-1,1\}^n}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z})]$

Next:

$$\mathop{\mathbb{E}}_{x,y \in \{-1,1\}^n}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z})]$$

By Fourier Expansion Theorem

$$= \mathop{\mathbb{E}}_{x,y \in \{-1,1\}^n}\left[\left(\sum_{S \subseteq [n]} \hat{f}(S)\chi_S(\boldsymbol{x})\right)\left(\sum_{T \subseteq [n]} \hat{f}(T)\chi_T(\boldsymbol{y})\right)\left(\sum_{U \subseteq [n]} \hat{f}(U)\chi_U(\boldsymbol{z})\right)\right]$$

Distributing out the product of sums

$$= \mathop{\mathbb{E}}_{x,y \in \{-1,1\}^n}\left[\left(\sum_{S,T,U \subseteq [n]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})\right)\right]$$

By linearity of expectation

$$= \sum_{S,T,U \subseteq [n]} \hat{f}(S)\hat{f}(T)\hat{f}(U) \mathop{\mathbb{E}}_{x,y \in \{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$$

Proof of Sum-Of-Cubes Lemma (Continued)

$$\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\text{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U) \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$$

Claim. $\mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$ is 1 if $S = T = U$ and 0 otherwise. ✓

- Let $S\Delta T$ denote symmetric difference of sets S and T

$$\mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})] = \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\textstyle\prod_{i\in S} x_i \prod_{i\in T} y_i \prod_{i\in U} z_i]$$

$$= \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\textstyle\prod_{i\in S} x_i \prod_{i\in T} y_i \prod_{i\in U} x_i y_i]$$

Since $\mathbf{z} = \mathbf{x} \circ \mathbf{y}$

$$= \mathop{\mathbb{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\textstyle\prod_{i\in S\Delta U} x_i \prod_{i\in T\Delta U} y_i]$$

Since $x_i^2 = y_i^2 = 1$

$$= \mathop{\mathbb{E}}_{\mathbf{x}\in\{-1,1\}^n}[\textstyle\prod_{i\in S\Delta U} x_i] \cdot \mathop{\mathbb{E}}_{\mathbf{y}\in\{-1,1\}^n}[\textstyle\prod_{i\in S\Delta U} y_i]$$

Since $\mathbf{x}$ and $\mathbf{y}$ are independent

$$= \textstyle\prod_{i\in S\Delta U} \mathop{\mathbb{E}}_{\mathbf{x}\in\{-1,1\}^n}[x_i] \cdot \prod_{i\in T\Delta U} \mathop{\mathbb{E}}_{\mathbf{y}\in\{-1,1\}^n}[y_i]$$

Since $\mathbf{x}$ and $\mathbf{y}$'s coordinates are independent

$$= \textstyle\prod_{i\in S\Delta U} \mathop{\mathbb{E}}_{x_i\in\{-1,1\}}[x_i] \cdot \prod_{i\in T\Delta U} \mathop{\mathbb{E}}_{y_i\in\{-1,1\}}[y_i]$$

$$= \begin{cases} 1 & \text{when } S\Delta U = \emptyset \text{ and } T\Delta U = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Proof of Sum-Of-Cubes Lemma (Done)

$$\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2}+\frac{1}{2}\sum_{S,T,U\subseteq[n]}\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$$

$$= \frac{1}{2}+\frac{1}{2}\sum_{S\subseteq[n]}\hat{f}(S)^3$$

Sum-Of-Cubes Lemma. $\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2}+\frac{1}{2}\sum_{S\subseteq[n]}\hat{f}(S)^3$ ✔

Proof of Correctness Theorem

Correctness Theorem (restated)

If f is ε-far from linear then $\Pr[\mathrm{BLR}(f) \text{ accepts}] \le 1 - \varepsilon$.

Proof: Suppose to the contrary that

$$1 - \varepsilon < \Pr_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[\mathrm{BLR}(f) \text{accepts}]$$

$$= \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{f}(S)^3$$

By Sum-Of-Cubes Lemma

$$\le \frac{1}{2} + \frac{1}{2} \cdot \left(\max_{S \subseteq [n]} \hat{f}(S) \right) \cdot \sum_{S \subseteq [n]} \hat{f}(S)^2$$

Since $\hat{f}(S)^2 \ge 0$

$$= \frac{1}{2} + \frac{1}{2} \cdot \left(\max_{S \subseteq [n]} \hat{f}(S) \right)$$

Parseval Equality

- Then $\max_{S \subseteq [n]} \hat{f}(S) > 1 - 2\varepsilon$. That is, $\hat{f}(T) > 1 - 2\varepsilon$ for some $T \subseteq [n]$.
- But $\hat{f}(T) = \langle f, \chi_T \rangle = 1 - 2 \cdot$ (fraction of *disagreements* between f and χ_T)
- f disagrees with a linear function χ_T on $< \varepsilon$ fraction of values. ※

Summary

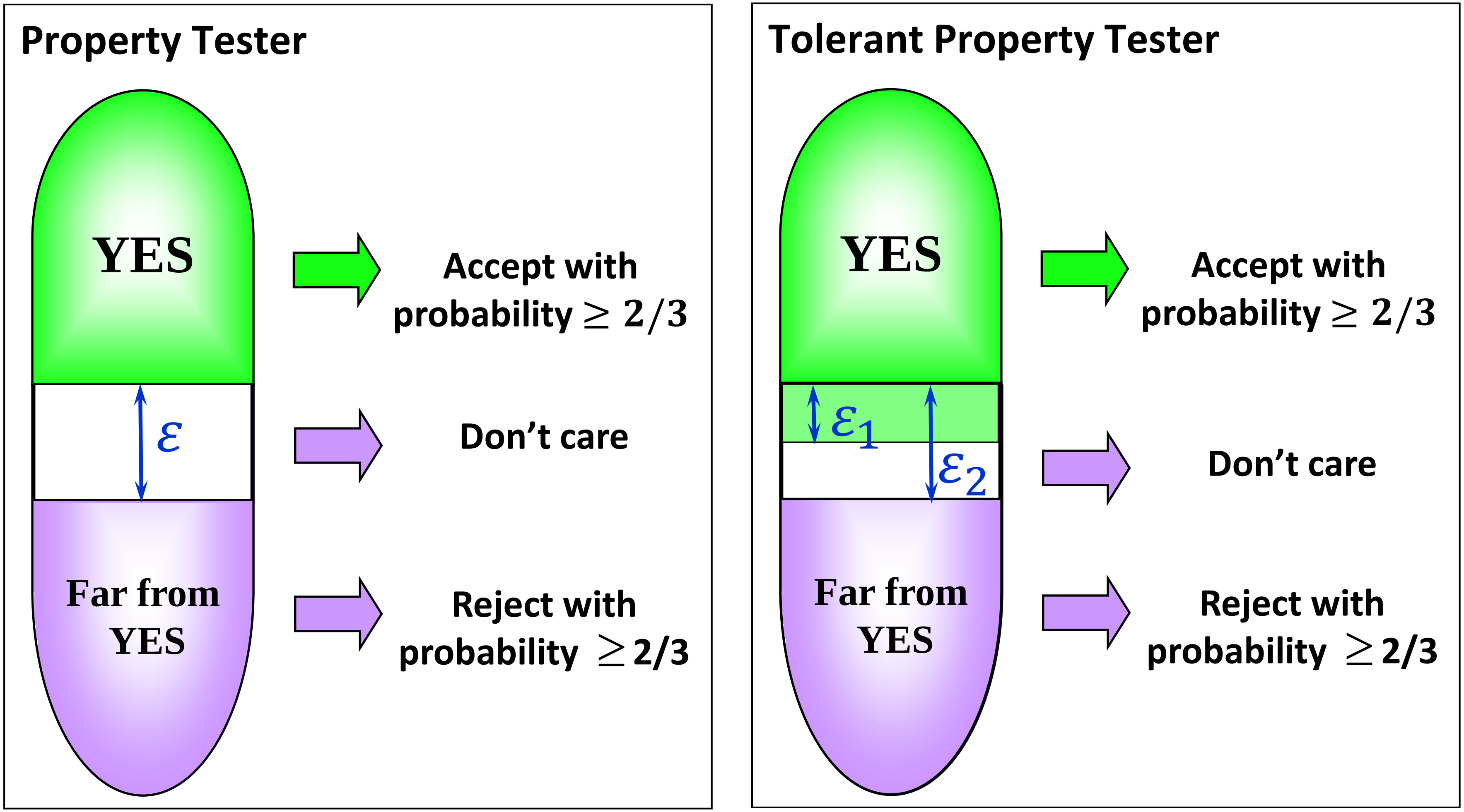
BLR tests whether a function $f: \{0,1\}^n \to \{0,1\}$ is

linear or ε-far from linear

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

in $O\left(\frac{1}{\varepsilon}\right)$ time.

Tolerant Property Testing [Parnas Ron Rubinfeld]



Two objects are at distance ε = they differ in an ε fraction of places

Equivalent problem: approximating distance to the property with additive error.

Distance Approximation to Property $\mathcal{P}$

Input: Parameter $\varepsilon \in (0,1/2]$ and query access to an object f

$$dist(f,\mathcal{P}) = \min_{g \in \mathcal{P}} dist(f,g)$$

$$dist(f,g) = \text{fraction of representation on which } f \text{ and } g \text{ differ}$$

Output: An estimate $\hat{\varepsilon}$ such that w.p. $\geq \frac{2}{3}$

$$|\hat{\varepsilon} - dist(f,\mathcal{P})| \leq \varepsilon$$

Approximating Distance to Monotonicity for 0/1 Sequences

Input: Parameter $\varepsilon \in (0,1/2]$ and

a list of n zeros and ones (equivalently, $f: [n] \to \{0,1\}$)

Question: How far is this list to being sorted?

(Equivalently, how far is f from monotone?)

$\text{dist}(f, MONO)$ =distance from f to monotone

$\text{Dist}(f, MONO) = n \cdot \text{dist}(f, MONO)$

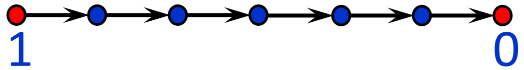
Note: $\text{Dist}(f, MONO) = n - |LIS|$,
where LIS is the longest increasing subsequence

Output: An estimate $\hat{\varepsilon}$ such that w.p. $\geq \frac{2}{3}$

$$|\hat{\varepsilon} - \text{dist}(f, MONO)| \leq \varepsilon$$

Today: can answer in $O\left(\frac{1}{\varepsilon^2}\right)$ time [Berman Raskhodnikova Yaroslavtsev]

Distance to Monotonicity over POset Domains

- Let f be a function over a partially ordered domain D.
- Violated pair: 

- The violation graph G_f is a directed graph with vertex set D whose edge set is the set of pairs (x, y) violated by f.
- VC_f is a minimum vertex cover of G_f
- MM_f is a maximum matching in G_f

Characterization of $Dist(f, \text{Mono})$ for $f: D \rightarrow \{0,1\}$ [FLNRRS 02]

$$Dist(f, \text{Mono}) = |\text{MM}_f| = |VC_f|$$

Distance to Monotonicity for 0/1 Sequences

- Let $f : [n] \to \{0,1\}$
- Great notation switch: $g_i = (-1)^{f(i)}$ for $i \in [n]$
- Cumulative sums: $s_0 = 0$ and $s_i = s_{i-1} + g_i$ for $i \in [n]$
- Final sum: $s_f = s_n$
- Maximum sum: $m_f = \max_{i=0}^{n} s_i$

dist$(f, Mono)$ for $f : [n] \to \{0,1\}$ [Berman Raskhodnikova Yaroslavtsev]

$$Dist(f, \mathrm{Mono}) = \frac{n - 2m_f + s_f}{2}$$

Proof:

1. Construct a matching of that size
2. Construct a vertex cover of that size.

Distance to Monotonicity for 0/1 Sequences

Characterization dist$(f, Mono)$ for $f: [n] \to \{0,1\}$

$$Dist(f, \text{Mono}) = \frac{n - 2m_f + s_f}{2}$$

Proof: (1) Construct a matching that leaves $2m_f - s_f$ nodes unmatched

Idea: For each edge chosen for the matching, perform operations on vector g that make it shorter while the maximum and the final sum remain unchanged.

While there exists an index i such that $g_i = -1$ and $g_{i+1} = 1$

match the vertices that contributed g_i and g_{i+1};

remove g_i and g_{i+1} from g.

- Let k be the length of the sequence after this procedure halted.
- Then g consists of 1's followed by -1's.
- $s_f = m_f - (k - m_f)$
- $k = 2m_f - s_f$
- The construction matches $n - k = n - 2m_f + s_f$ vertices.

The number of 1's is m_f.

Distance to Monotonicity for 0/1 Sequences

Characterization dist$(f, Mono)$ for $f: [n] \to \{0,1\}$

$$Dist(f, \mathrm{Mono}) = \frac{n - 2m_f + s_f}{2}$$

Proof: (2) Construct a vertex cover.

Idea: Consider the edges of the matching we constructed in the opposite order of insertion.

Distance to Monotonicity: Algorithm

Algorithm (**Input:** ε, n; query $acess\ to\ f: [n] \rightarrow \{0,1\}$)

1. Sample a random subset $\boldsymbol{S} \subset [n]$
   where each element is included w.p. s/n independently
2. Let $\tilde{f} = f_{|S}$
3. Compute $\tilde{\varepsilon} = Dist(\tilde{f}, Mono)/s$
4. **Return** $\tilde{\varepsilon}$

- Let $\varepsilon_f = dist(f, Mono) = Dist(f, Mono)/n$

Theorem

$$\varepsilon_f - \sqrt{2\varepsilon_f/s} \leq \mathbb{E}[\tilde{\varepsilon}] \leq \varepsilon_f$$

$$\text{Var}[\tilde{\varepsilon}] = O(\varepsilon_f/s)$$

Proof idea: Let $Z(S) = Dist(\tilde{f}, Mono)$

We'll define random variables $X(S)$ and $\mathrm{Y}(S)$, such that $X(S) \leq Z(S) \leq Y(S)$

$X(S)$ will be in terms of matching MM_f; $Y(S)$ in terms of vertex cover VC_f