Lower Bounds for Testing Properties of Functions on Hypergrids

Grigory Yaroslavtsev

http://grigory.us

Joint with:

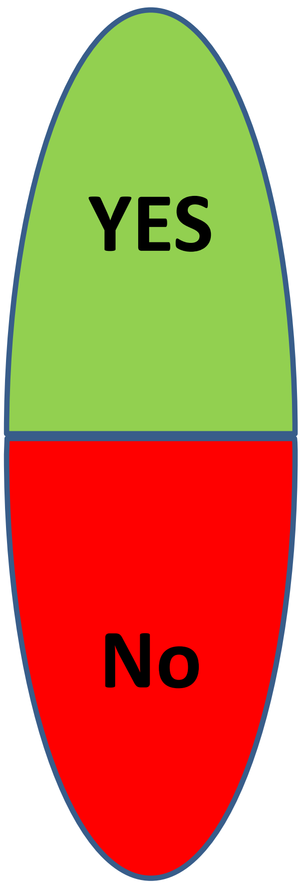
Eric Blais (MIT)

Sofya Raskhodnikova (PSU)

Property Testing

[Goldreich, Goldwasser, Ron, Rubinfeld, Sudan]

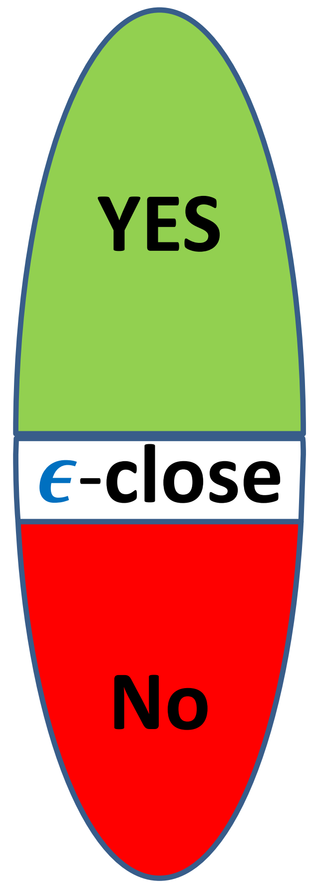
Randomized algorithm



$\Rightarrow$ **Accept with probability** $\geq \frac{2}{3}$

$\Rightarrow$ **Reject with probability** $\geq \frac{2}{3}$

Property tester



$\Rightarrow$ **Accept with probability** $\geq \frac{2}{3}$

$\Rightarrow$ **Don’t care**

$\Rightarrow$ **Reject with probability** $\geq \frac{2}{3}$

ϵ**-close :** $\leq \epsilon$ fraction can be changed to become **YES**

Ultra-fast Approximate Decision Making

Property Testing

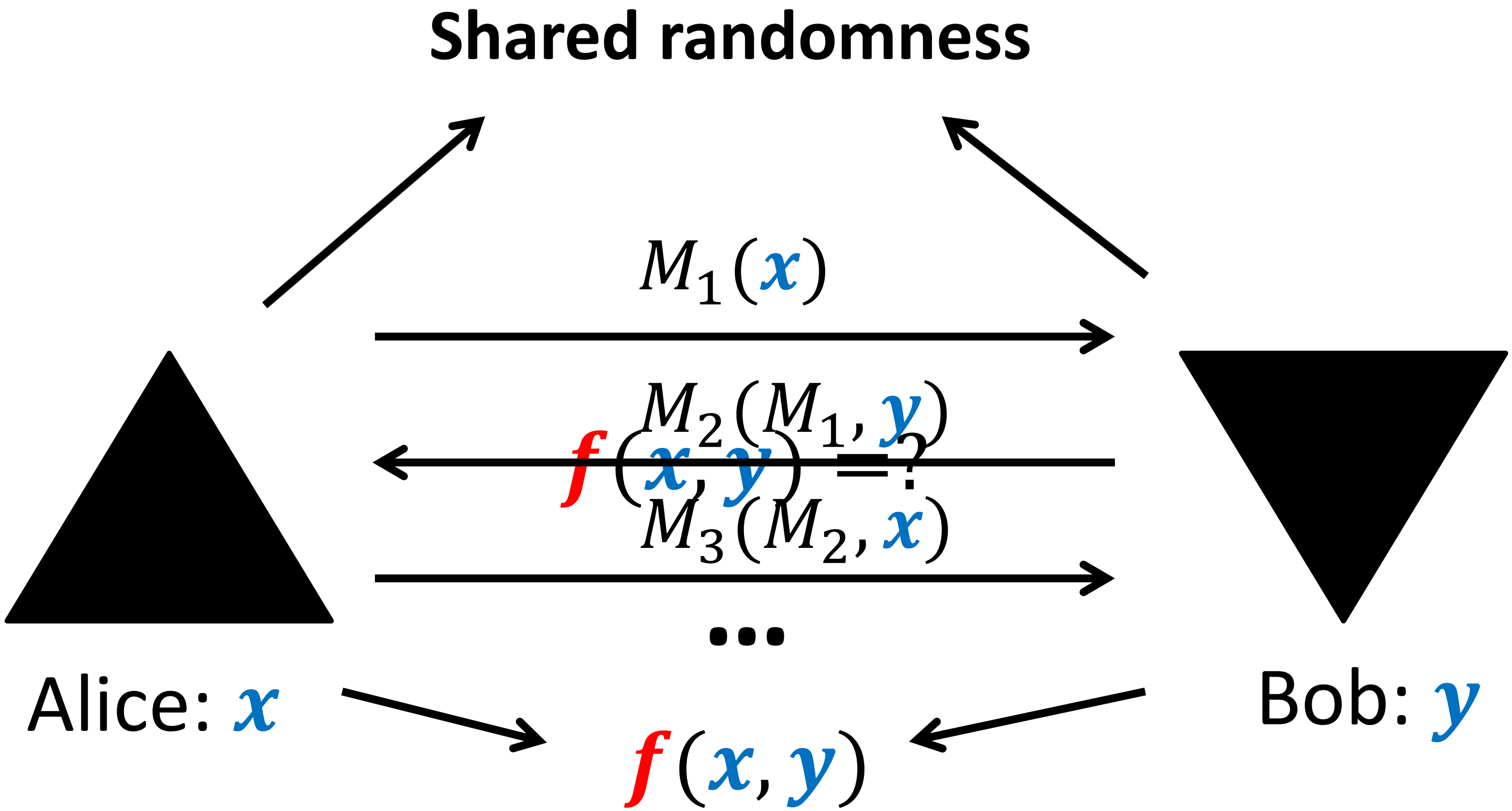
[Goldreich, Goldwasser, Ron, Rubinfeld, Sudan]

Property $\boldsymbol{P}$ = set of **YES** instances

Query complexity of testing $\boldsymbol{P}$:

- $\boldsymbol{Q}_{\epsilon}(\boldsymbol{P})$ = Adaptive queries
- $\boldsymbol{Q}_{\epsilon}^{na}(\boldsymbol{P})$ = Non-adaptive (all queries at once)
- $\boldsymbol{Q}_{\epsilon}^{r}(\boldsymbol{P})$ = Queries in r rounds ($\boldsymbol{Q}_{\epsilon}^{na}(\boldsymbol{P}) = \boldsymbol{Q}_{\epsilon}^{1}(\boldsymbol{P})$)

Communication Complexity [Yao'79]



- $R(f)$ = min. communication (error 1/3)
- $R^k(f) =$ min. k-round communication (error 1/3)

$k/2$-disjointness ⇒ k-linearity

[Blais, Brody,Matulef'11]

- k-linear function: $\{0,1\}^n \to \{0,1\}$

$$\bigoplus_{i \in S} x_i = x_{i_1} \oplus x_{i_2} \oplus \cdots \oplus x_{i_k}$$

where $|S| = k$

- $k/2$-Disjointness: $S, T \subseteq [n], |S| = |T| = \frac{k}{2}$

$$f(S, T) = 1, \text{ iff } |S \cap T| = 0.$$

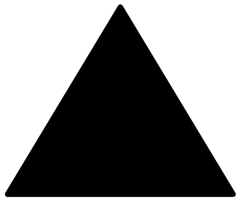
f: $|S \cap T| = 0$?

Alice:
$S \subseteq [n], |S| = k/2$

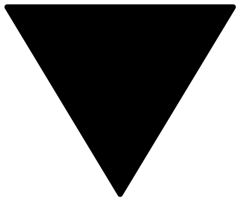
Bob:
$T \subseteq [n], |T| = k/2$

k/2-disjointness ⇒ k-linearity

[Blais, Brody, Matulef'11]



$$\chi = \chi_S \oplus \chi_T$$



$S \subseteq [n], |S| = k/2$

$\chi_S = \bigoplus_{i \in S} x_i$

$T \subseteq [n], |T| = k/2$

$\chi_T = \bigoplus_{i \in T} x_i$

- $S \cap T = \emptyset \Rightarrow \chi$ is k-linear
- $S \cap T \neq \emptyset \Rightarrow \chi$ is $(< k)$-linear, ½-far from k-linear

- Test χ for k-linearity using shared randomness
- To evaluate $\chi(x)$ exchange $\chi_S(x)$ and $\chi_T(x)$ (2 bits)
- $R\left(\frac{k}{2}\text{–Disjointness}\right) \leq 2 \cdot Q_{1/2}(k\text{–Linearity})$

$\boldsymbol{k}$-Disjointness

- $\boldsymbol{R}(\boldsymbol{k}\text{–Disjointness}) = \Theta(\boldsymbol{k})$ [Razborov, Hastad-Wigderson]
- $\boldsymbol{R}^{\mathbf{1}}(\boldsymbol{k}\text{–Disjointness}) = \Theta(\boldsymbol{k} \log \boldsymbol{k})$

[Folklore + Dasgupta, Kumar, Sivakumar'12; Buhrman, Garcia-Soriano, Matsliah, De Wolf'12]

- $\boldsymbol{R}^{\boldsymbol{r}}(\boldsymbol{k}\text{–Disjointness}) = \Theta(\boldsymbol{k}\ \mathrm{ilog}^r \boldsymbol{k})$,

where $ilog^r k = \log \underbrace{\log \ldots \log}_{r \text{ times}} k$ [Saglam, Tardos'13]

$$\Omega(\boldsymbol{k}\ \mathrm{ilog}^r \boldsymbol{k}) = \boldsymbol{Q}^{\boldsymbol{r}}{}_{\mathbf{1/2}}(\boldsymbol{k}\text{–Linearity})$$

- $\boldsymbol{R}(\boldsymbol{k}\text{–Disjointness}) = \boldsymbol{\alpha}\boldsymbol{k} + o(\boldsymbol{k})$[Braverman, Garg, Pankratov, Weinstein'13]

Property testing lower bounds via CC

- Monotonicity, Juntas, Low Fourier degree, Small Decision Trees [Blais, Brody, Matulef'11]
- Small-width OBDD properties [Brody, Matulef, Wu'11]
- Lipschitz property [Jha, Raskhodnikova'11]
- Codes [Goldreich'13, Gur, Rothblum'13]
- Number of relevant variables [Ron, Tsur'13]

(Almost) all: Boolean functions over Boolean hypercube

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

$M_{\boldsymbol{m},\boldsymbol{n}}$ = monotone functions over $[\boldsymbol{m}]^{\boldsymbol{n}}$

$$\boldsymbol{Q}^1(M_{\boldsymbol{m},\boldsymbol{n}}) = \Omega(\boldsymbol{n} \log \boldsymbol{m})$$

Previous for monotonicity on the line ($\boldsymbol{n} = 1$):

- $\boldsymbol{Q}^1(M_{\boldsymbol{m},\mathbf{1}}) = \Theta(\log \boldsymbol{m})$ [Ergun, Kannan, Kumar, Rubinfeld, Viswanathan'00]
- $\boldsymbol{Q}(M_{\boldsymbol{m},\mathbf{1}}) = \Omega(\log \boldsymbol{m})$ [Fischer'04]

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

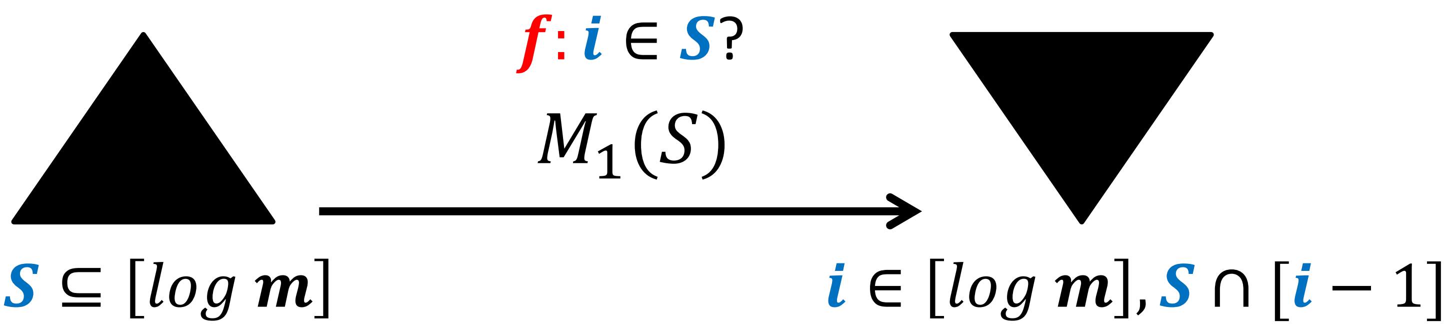
- **Proof ideas:**
  - Reduction from Augmented Index (widely used in streaming, e.g [Jayram, Woodruff'11; Molinaro, Woodruff, Y.'13])
  - Fourier analysis over $\{0,1\}^n$ basis of characters => Fourier analysis over $[\boldsymbol{m}]^{\boldsymbol{n}}$: basis of Walsh functions

- **Case $\boldsymbol{n} = \mathbf{1}$**: Any non-adaptive tester for monotonicity of $f: [\boldsymbol{m}] \to [r]$ has complexity

$$\Omega(\min(\log \mathbf{m}, \log \mathbf{r}))$$

Functions $[\boldsymbol{m}] \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

- **Augmented Index**: $\boldsymbol{S}$; $(\boldsymbol{i}, \boldsymbol{S} \cap [\boldsymbol{i}-1])$



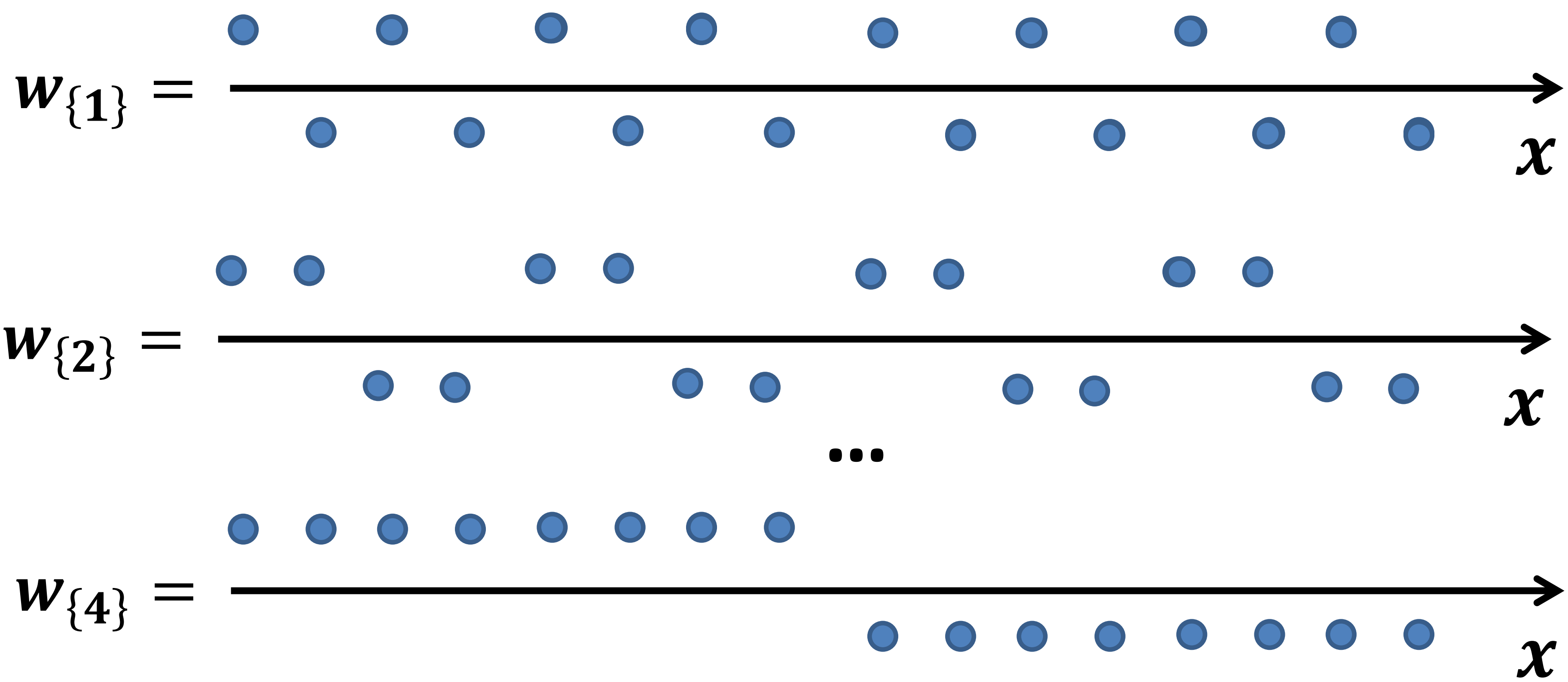
- $\boldsymbol{R}^1[\text{Augmented Index}] = \Omega(|\boldsymbol{S}|)$ [Miltersen, Nisan, Safra, Wigderson, 98]

Functions $[\boldsymbol{m}] \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

Walsh functions: For $S \subseteq [\log \boldsymbol{m}]$, $w_S : [m] \to \{-1,1\}$:

$$w_S(x) = \prod_{i \in S} (-1)^{x_i},$$

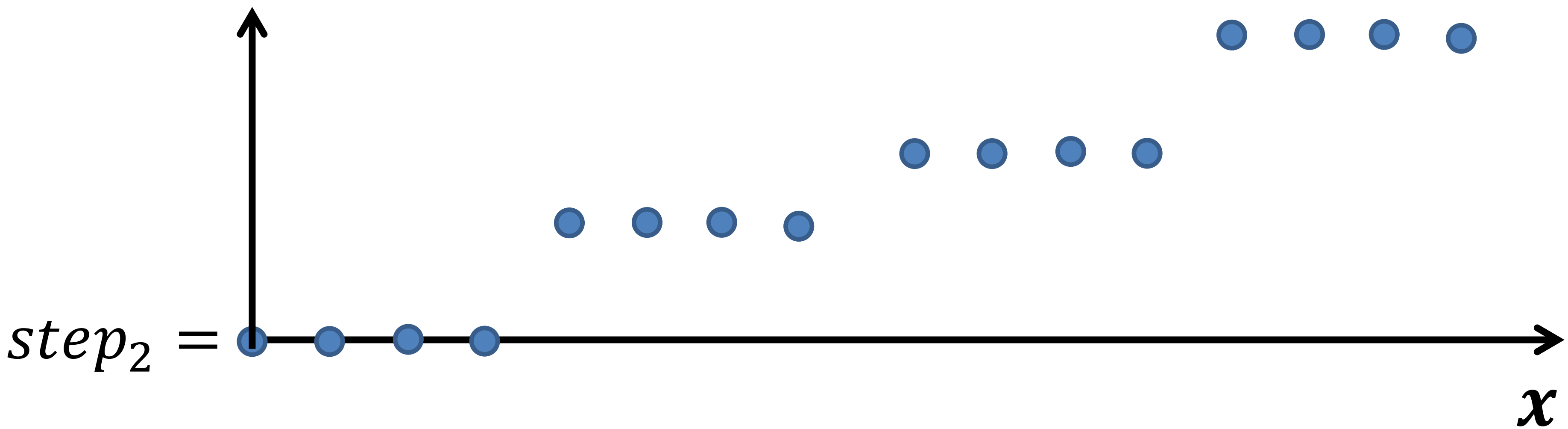
where x_i is the i-th bit of x.



Functions $[\boldsymbol{m}] \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

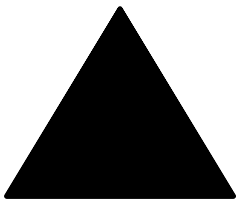
Step functions. For $i \in [\log \boldsymbol{m}]$: $step_i \colon [\boldsymbol{m}] \to \left[\frac{\boldsymbol{m}}{2^i}\right]$:

$$step_i(x) = \lceil x/2^i \rceil$$



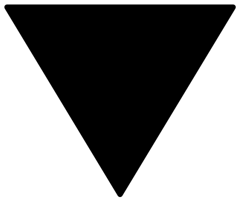
Functions $[\boldsymbol{m}] \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

- **Augmented Index** $\Rightarrow$ Monotonicity Testing



$\boldsymbol{S} \subseteq [log\ \boldsymbol{m}]$

$$\chi = w_{\boldsymbol{S} \cap [\boldsymbol{i}, \ldots, log\ \boldsymbol{m}]} + 2\ step_{\boldsymbol{i}}$$
$$= w_{\boldsymbol{S}} \oplus w_{\boldsymbol{S} \cap [\boldsymbol{i}-1]} + 2\ step_{\boldsymbol{i}}$$



$\boldsymbol{i} \in [log\ \boldsymbol{m}], \boldsymbol{S} \cap [\boldsymbol{i} - 1]$

- $\boldsymbol{i} \notin \boldsymbol{S} \Rightarrow \chi$ is monotone
- $\boldsymbol{i} \in \boldsymbol{S} \Rightarrow \chi$ is ¼ -far from monotone
- Only $\boldsymbol{i}$-th frequency matters: higher frequencies are cancelled, lower don't affect monotonicity
- Thus, $\boldsymbol{Q}^{\mathbf{1}}(M_{\boldsymbol{m},\mathbf{1}}) = \Omega(\log \boldsymbol{m})$

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

▲

$\boldsymbol{S} \subseteq [\boldsymbol{n}\ log\ \boldsymbol{m}]$

$\Downarrow$

$\boldsymbol{S}_{\mathbf{1}}, \dots, \boldsymbol{S}_{\boldsymbol{n}} \subseteq [\log \boldsymbol{m}]$

▼

$\boldsymbol{i} \in [\boldsymbol{n}\ log\ \boldsymbol{m}], \boldsymbol{S} \cap [\boldsymbol{i} - 1]$

$\Downarrow$

$(m, m, \dots, m, \boldsymbol{i}_{\boldsymbol{j}^*}, 0,0, \dots, 0)$

$\boldsymbol{S}_1, \dots, \boldsymbol{S}_{\boldsymbol{j}^*-1}, \boldsymbol{S}_{\boldsymbol{j}^*} \cap [\boldsymbol{i}_{\boldsymbol{j}^*} - 1]$

Embed into $\boldsymbol{j}^*$-th coordiante using $\boldsymbol{n}$-dimensional Walsh and step functions:

- Walsh functions: $\boldsymbol{w}_{\boldsymbol{S}}(x_1, \dots, x_{\boldsymbol{n}}) = \prod_{j=1}^{n} w_{\boldsymbol{S}_j}(x_j)$
- Step functions: $\mathbf{step}_{\boldsymbol{i}}(x_1, \dots, x_{\boldsymbol{n}}) = \sum_{j=1}^{n} step_j(x_j)$

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

▲

$\boldsymbol{S}_{\boldsymbol{1}}, \dots, \boldsymbol{S}_{\boldsymbol{n}} \subseteq [\log \boldsymbol{m}]$

▼

$(m, m, \dots, m, \boldsymbol{i}_{\boldsymbol{j}^*}, 0,0, \dots, 0)$

$\boldsymbol{S}_{1,\dots,} \boldsymbol{S}_{\boldsymbol{j}^*-1}, \boldsymbol{S}_{\boldsymbol{j}^*} \cap [\boldsymbol{i}_{\boldsymbol{j}^*} - 1]$

$$\boldsymbol{\chi}(\boldsymbol{x}_{\boldsymbol{1}}, \dots, \boldsymbol{x}_{\boldsymbol{n}}) =$$

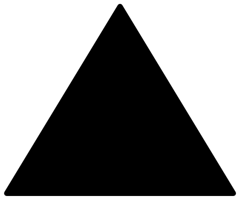
$$w_{\boldsymbol{S}_{\boldsymbol{j}^*} \cap \left[\boldsymbol{i}_{\boldsymbol{j}^*}, \dots, \log \boldsymbol{m}\right]}(x_{\boldsymbol{j}^*}) \oplus \prod_{j=\boldsymbol{j}^*+1}^{\boldsymbol{n}} w_{\boldsymbol{S}_j}(x_j) + 2\,\mathbf{step}_{\boldsymbol{i}}(x_1, .., x_{\boldsymbol{n}}) =$$

$$\boldsymbol{w}_{\boldsymbol{S}}(x_1, .., x_n) \oplus \prod_{j=1}^{\boldsymbol{j}^*-1} w_{\boldsymbol{S}_j}(x_j) +$$

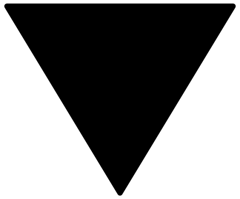
$$2\sum_{j=1}^{\boldsymbol{j}^*-1} \text{step}_m(x_j) + 2 step_{i_{\boldsymbol{j}^*}}(x_{\boldsymbol{j}^*}) + 2\sum_{j=\boldsymbol{j}^*+1}^{\boldsymbol{n}} step_0(x_j)$$

- Walsh functions: $\boldsymbol{w}_{\boldsymbol{S}}(x_1, \dots, x_{\boldsymbol{n}}) = \prod_{j=1}^{n} w_{\boldsymbol{S}_j}(x_j)$
- Step functions: $\mathbf{step}_{\boldsymbol{i}}(x_1, \dots, x_{\boldsymbol{n}}) = \sum_{j=1}^{n} step_j(x_j)$

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]



$\boldsymbol{S_1}, \ldots, \boldsymbol{S_n} \subseteq [\log \boldsymbol{m}]$



$(m, m, \ldots, m, \boldsymbol{i_{j^*}}, 0,0, \ldots, 0)$

$\boldsymbol{S}_1, \ldots, \boldsymbol{S}_{\boldsymbol{j}^*-1}, \boldsymbol{S}_{\boldsymbol{j}^*} \cap [\boldsymbol{i}_{\boldsymbol{j}^*} - 1]$

$\chi(\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}) =$

$\boldsymbol{w_S}(x_1, \ldots, x_n) \oplus \prod_{j=1}^{\boldsymbol{j}^*-1} w_{\boldsymbol{S}_j}(x_j) + 2step_{i_{\boldsymbol{j}^*}}(x_{\boldsymbol{j}^*}) + 2\sum_{j=\boldsymbol{j}^*+1}^{\boldsymbol{n}} x_j$

- Only coordinate $\boldsymbol{j}^*$ matters:
  - Coordinates < $\boldsymbol{j}^*$ cancelled by Bob's Walsh terms
  - Coordinates > $\boldsymbol{j}^*$ cancelled by Bob's Step terms
  - Coordinate $\boldsymbol{j}^*$ behaves as in the $\boldsymbol{n} = 1$ case

Functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$ [Blais, Raskhodnikova, Y.]

- $M_{\boldsymbol{m},\boldsymbol{n}}$ = monotone functions over $[\boldsymbol{m}]^{\boldsymbol{n}}$

$$\boldsymbol{Q}^1(M_{\boldsymbol{m},\boldsymbol{n}}) = \Omega(\boldsymbol{n} \log \boldsymbol{m})$$

- $L_{\boldsymbol{m},\boldsymbol{n}}$ = c-Lipschitz functions over $[\boldsymbol{m}]^{\boldsymbol{n}}$
- $C^s_{\boldsymbol{m},\boldsymbol{n}}$ = separately convex functions over $[\boldsymbol{m}]^{\boldsymbol{n}}$
- $M^k_{\boldsymbol{m},\boldsymbol{n}}$ = monotone axis-parallel k-th derivative over $[\boldsymbol{m}]^{\boldsymbol{n}}$
- $C_{\boldsymbol{m},\boldsymbol{n}}$ = convex functions over $[\boldsymbol{m}]^{\boldsymbol{n}}$
  - Can't be expressed as a property of axis-parallel derivatives!

Thm. [BRY] For all these properties $\boldsymbol{Q}^1 = \Omega(\boldsymbol{n} \log \boldsymbol{m})$

These bounds are optimal for $M_{\boldsymbol{m},\boldsymbol{n}}$ and $L_{\boldsymbol{m},\boldsymbol{n}}$ [Chakrabarty, Seshadhri, '13]

Open Problems

- Adaptive bounds and round vs. query complexity tradeoffs for functions $[\boldsymbol{m}]^{\boldsymbol{n}} \to \mathbb{R}$
  - Only known: $\boldsymbol{Q}\left(M_{\boldsymbol{m},\boldsymbol{n}}\right) = \Omega(\boldsymbol{n} \log \boldsymbol{m})$ [Fischer'04; Chakrabarty Seshadhri'13]
- Inspired by connections of CC and Information Complexity
  - Direct information-theoretic proofs?
  - Round vs. query complexity tradeoffs in property testing?
- Testing functions $[\mathbf{0}, \mathbf{1}]^{\boldsymbol{n}} \to \mathbb{R}$
  - L_p-testing model [Berman, Raskhodnikova, Y. '14]
  - Testing convexity: $2^{O(\boldsymbol{n} \, log \, \boldsymbol{n})}$ vs. $\Omega(\boldsymbol{n})$?