Strong Direct Sum for Randomized Query Complexity

Eric Blais
University of Waterloo

Joshua Brody
Swarthmore College

Conference on Computational Complexity
New Brunswick, New Jersey
July 18, 2019

Outline

- **Introduction**
- Strong Direct Sum
- Query Resistance
- Separation Theorem
- Open Problems

Direct Sum Theorems

Does computing f(x) on k copies scale with k?

Direct Sum Theorems

Does computing f(x) on k copies scale with k?

Direct Sum Theorem: Computing **k** copies of **f** requires **k** times the resources

Direct Product Theorem: Success prob. of computing **k** copies of **f** with **<< k** resources is $2^{-\Omega(k)}$

Direct Sum Theorems

Does computing f(x) on k copies scale with k?

Direct Sum Theorem: Computing **k** copies of **f** requires **k** times the resources

Direct Product Theorem: Success prob. of computing **k** copies of **f** with **<< k** resources is $2^{-\Omega(k)}$

Strong Direct Sum: computing k copies of f w/error *ε* requires >> k times the resources

Our Main Results

Strong Direct Sum for *average* query complexity:

For any f and any k, computing f^k satisfies:

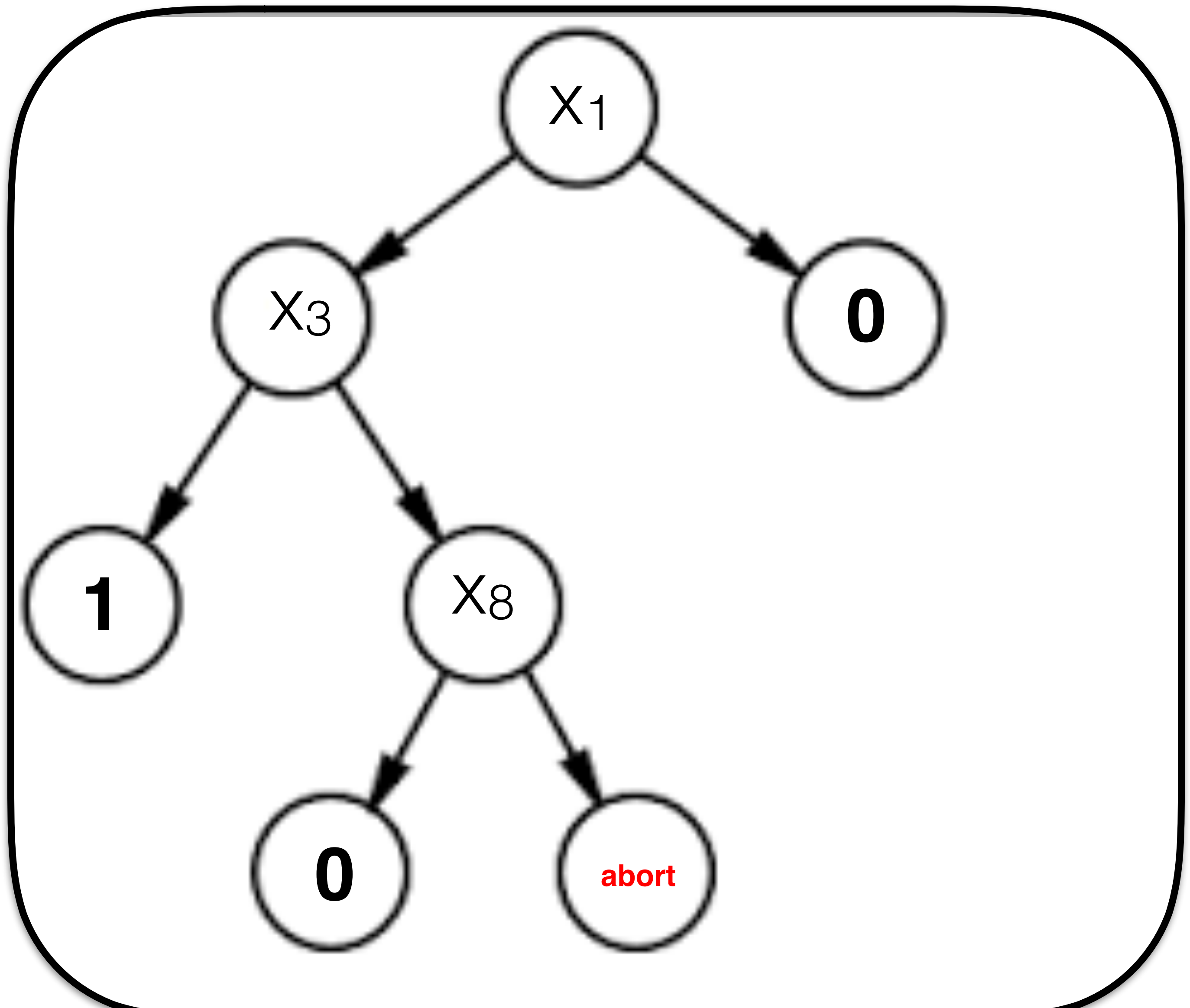
$$\bar{R}_{\varepsilon}(f^k) = \Theta(k\bar{R}_{\varepsilon/k}(f))$$

Separation Theorem: for all $\varepsilon > 2^{-n^{1/3}}$, there is total function $f : \{0,1\}^N \rightarrow \{0,1\}$ such that $\bar{R}_{\varepsilon}(f) = \Theta(R(f)\log(1/\varepsilon))$

Corollary: There is f such that $R_{\varepsilon}(f^k) = \Theta(k\log(k)R_{\varepsilon}(f))$

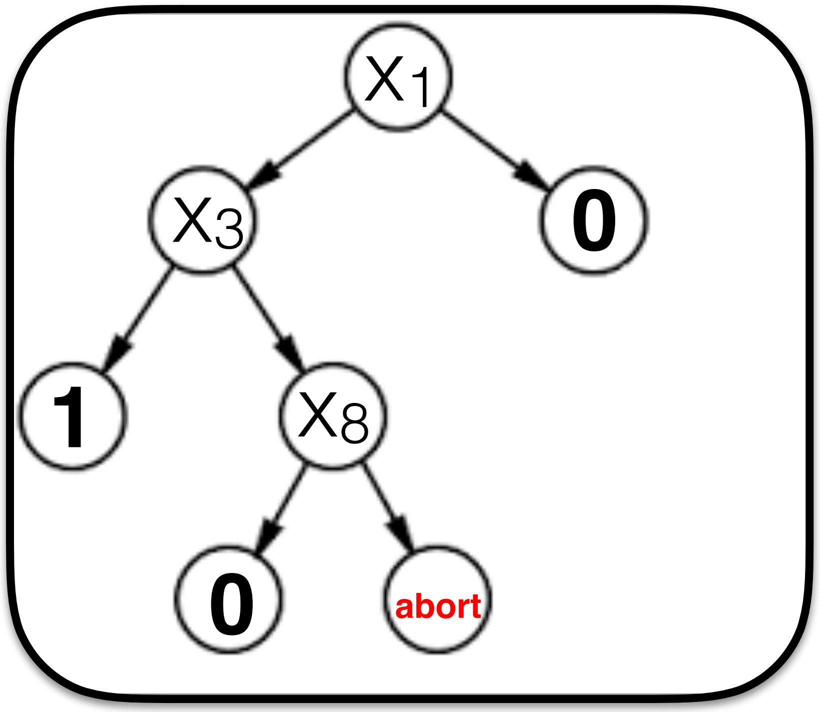
Query Complexity

aka Decision Tree Complexity



Query Complexity

aka Decision Tree Complexity



Decision Tree for $f: \{0,1\}^n \to \{0,1\}$:

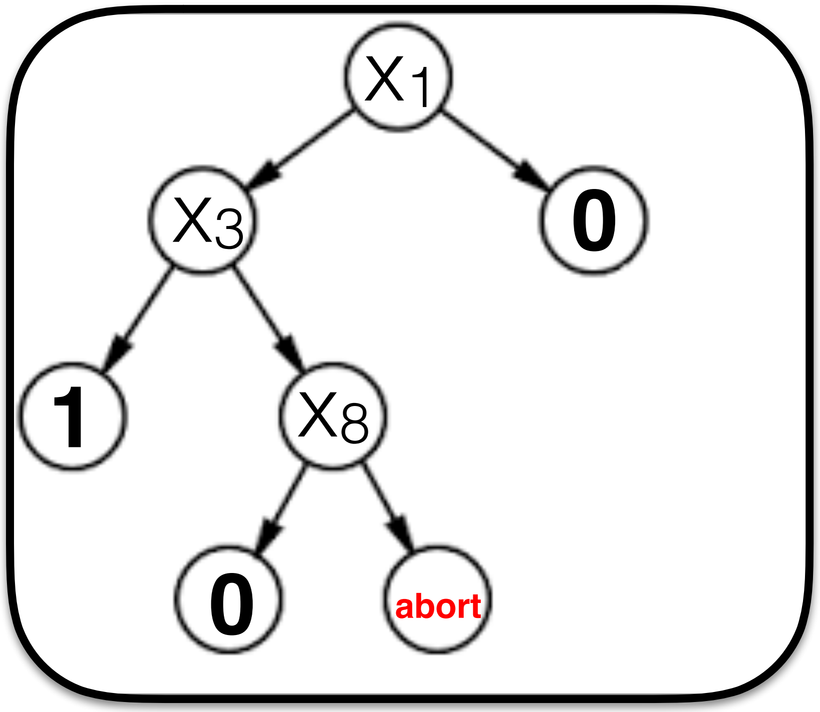
- internal nodes labeled w/input bits x_i
- leaves labeled w/output or **ABORT**
- **cost(T,x):** depth of **T** on input x

Randomized DT: distribution *A* on decision trees

- $\mathrm{cost}(A) = \max_{T,x} \mathrm{cost}(T,x)$
- $\mathrm{acost}(A) = \max_x E_{T\sim A} [\mathrm{cost}(T, x)]$

Query Complexity

aka Decision Tree Complexity



Decision Tree for $f: \{0,1\}^n \to \{0,1\}$:

- internal nodes labeled w/input bits x_i
- leaves labeled w/output or **ABORT**
- **cost(T,x):** depth of **T** on input x

Randomized DT: distribution A on decision trees

- $\mathrm{cost}(A) = \max_{T,x} \mathrm{cost}(T,x)$
- $\mathrm{acost}(A) = \max_x E_{T\sim A} [\mathrm{cost}(T, x)]$

Distributional QC $D^{\mu}_{\delta,\varepsilon}(f)$**:** min $E_x[\mathrm{cost}(T,x)]$ s.t. $\Pr[\text{abort}] \leq \delta$ and $\Pr[\text{error}] \leq \varepsilon$

Randomized QC $R_{\delta,\varepsilon}(f)$**:** minimum cost of randomized algorithm s.t.

$\Pr[\text{abort}] \leq \delta$ and $\Pr[\text{error}] \leq \varepsilon$

Average case Randomized QC $\overline{R}_{\varepsilon}(f)$ **:**

minimum acost of randomized algorithm s.t. $\Pr[\text{error}] \leq \varepsilon$

Basic Results

Minimax Lemma: $\max_\mu D^{\mu}_{2\delta,2\varepsilon}(f) \leq R_{\delta,\varepsilon}(f) \leq \max_\mu D^{\mu}_{\delta/2,\varepsilon/2}(f)$

Error Reduction: $R_{O(1/t),\, O(1/t)}(f) \leq O(\log(t) R_{1/2,\, 1/3}(f))$

Average QC vs Aborts: $\delta R_{\delta,\varepsilon}(f) \leq \bar{R}_{\varepsilon}(f) \leq R_{\delta,(1-\delta)\varepsilon}(f)/(1-\delta)$

Basic Results

Average QC vs Aborts: $\delta R_{\delta,\varepsilon}(f) \le \bar{R}_{\varepsilon}(f) \le R_{\delta,(1-\delta)\varepsilon}(f)/(1-\delta)$

First inequality:

Algorithm ***A***: ***ε***-error,

acost(A)= q

Second inequality:

Algorithm ***B'***: ***(1-δ)ε***-error,

δ-abort, **q** queries.

Basic Results

Average QC vs Aborts: $\delta R_{\delta,\varepsilon}(f) \le \bar{R}_{\varepsilon}(f) \le R_{\delta,(1-\delta)\varepsilon}(f)/(1-\delta)$

First inequality:

Algorithm ***A***: ***ε***-error,

acost(A)= q

```
Algorithm B(x) {
  emulate A(x)
  abort if > q/δ queries
}
```

Second inequality:

Algorithm ***B'***: ***(1-δ)ε***-error,

δ-abort, **q** queries.

```
Algorithm A'(x) {
  repeat:
    emulate B'(x)
  until no aborts
}
```

Previous Work

Information Complexity: **[MWY13, MWY15]**

- strong direct sum for *information complexity w/aborts + error*
- applications for streaming/sketching algorithms

Direct Product Theorem: **[Drucker 12]**

- direct product theorems for randomized query complexity

Separation Theorems: **[GPW15, ABBLSS17]**

- query complexity separations based on *pointer functions*
- polynomial separation $\mathbf{R_0(f)}$ vs $\mathbf{R_\varepsilon(f)}$

Direct Sum Theorems:

- **[Jain Klauck Santha 10]:** $\mathbf{R}_\varepsilon(f^k) \geq \delta^2 k \mathbf{R}_{\varepsilon/(1-\delta)+\delta}(f)$
- **[Ben-David Kothari 18]:** $\overline{\mathbf{R}}_\varepsilon(f^k) \geq k\overline{\mathbf{R}}_\varepsilon(f)$

Our Results

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Separation Theorem: There is $f : \{0,1\}^N \rightarrow \{0,1\}$ such that for all $\varepsilon > 2^{-N^{1/3}}$, we have $R_{\delta,\varepsilon}(f) = \Omega(R_{1/3}(f)\log(1/\varepsilon))$

Corollary: There is f such that $R_{1/3}(f^k) = \Omega(k\log(k)R_{\varepsilon}(f))$

Our Results

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Separation Theorem: There is $f : \{0,1\}^N \rightarrow \{0,1\}$ such that for all $\varepsilon > 2^{-N^{1/3}}$, we have $R_{\delta,\varepsilon}(f) = \Omega(R_{1/3}(f)\log(1/\varepsilon))$

Corollary: There is f such that $R_{1/3}(f^k) = \Omega(k\log(k)R_\varepsilon(f))$

proof: $R_{1/3}(f^k) \geq R_{0,1/3}(f^k) = \Omega(kR_{1/5,40/3k}(f)) = \Omega(k\log(k)R_{1/3}(f))$

Our Results

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Separation Theorem: There is $f : \{0,1\}^N \to \{0,1\}$ such that for all $\varepsilon > 2^{-N^{1/3}}$, we have $R_{\delta,\varepsilon}(f) = \Omega(R_{1/3}(f)\log(1/\varepsilon))$

Corollary: There is f such that $R_{1/3}(f^k) = \Omega(k\log(k)R_{\varepsilon}(f))$

proof: $R_{1/3}(f^k) \geq R_{0,1/3}(f^k) = \Omega(kR_{1/5,40/3k}(f)) = \Omega(k\log(k)R_{1/3}(f))$

Key Technical result:

Query-resistant codes: probabilistic encoding $G: \Sigma \to \{0,1\}^N$ such that $N/3$ bits of $G(x)$ needed to learn anything about x

Outline

- Introduction
- **Strong Direct Sum**
- Query Resistance
- Separation Theorem
- Open Problems

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Let ***A*** be an **ε-error** algorithm for f^k with **q** queries.

Goal: ***(ε/k)-error*** algorithm ***B*** for ***f*** with **q/k** queries.

Let $\mathbf{y} = (y_1, \ldots, y_k)$.

Embed(y,i,x) := y, w/i-th coord replaced by **x**.

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Let ***A*** be an **ε-error** algorithm for f^k with **q** queries.

Goal: ***(ε/k)-error*** algorithm ***B*** for ***f*** with **q/k** queries.

Let $y = (y_1, \ldots, y_k)$.

Embed(y,i,x) := **y**, w/i-th coord replaced by **x**.

```
Algorithm B(x) {
  carefully select y,i
  emulate A(EMBED(y,i,x))
  abort if problems found
}
```

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

Let ***A*** be an **ε-error** algorithm for f^k with **q** queries.

Goal: ***(ε/k)-error*** algorithm ***B*** for ***f*** with **q/k** queries.

Let $y = (y_1,\ldots, y_k)$.

Embed(y,i,x) := y, w/i-th coord replaced by **x**.

```
Algorithm B(x) {
  carefully select y,i
  emulate A(EMBED(y,i,x))
  abort if problems found
}
```

Intuition: success on typical coordinate $\geq 1- 10\varepsilon/k$

else overall success $< (1- 10\varepsilon/k)^k < 1-\varepsilon$

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

$$1-\varepsilon \le \Pr_{Y\sim\mu^k}[A(Y) = f^k(Y)] = \prod_{i=1}^{k} \Pr_{Y\sim\mu^k}[A(Y)_i = f^k(Y)_i \mid A(Y)_{<i} = f^k(Y)_{<i}]$$

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

$$1-\varepsilon \le \Pr_{Y\sim\mu^k}[A(Y) = f^k(Y)] = \prod_{i=1}^{k} \Pr_{Y\sim\mu^k}[A(Y)_i = f^k(Y)_i \mid A(Y)_{<i} = f^k(Y)_{<i}]$$

Want: i such that

(1) conditional error very low:

$$\Pr[A \text{ err. on i-th coord.} \mid \text{correct on} < i] \le 10\ \varepsilon/k$$

(2) Expected # queries on i-th coord not too high:

$$E[\text{queries on i-th coord.}] \le 3q/k$$

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

$$1-\varepsilon \le \Pr_{Y\sim\mu^k}[A(Y) = f^k(Y)] = \prod_{i=1}^{k} \Pr_{Y\sim\mu^k}[A(Y)_i = f^k(Y)_i \mid A(Y)_{<i} = f^k(Y)_{<i}]$$

Want: i such that

(1) conditional error very low:

$$\Pr[\text{A err. on i-th coord.} \mid \text{correct on} < i] \le 10\,\varepsilon/k$$

(2) Expected # queries on i-th coord not too high:

$$E[\text{queries on i-th coord.}] \le 3q/k$$

Fact: at least ***2k/3*** coords. satisfy (1)

Fact: at least ***2k/3*** coords. satisfy (2)

$\Longrightarrow$ There is ***i**** satisfying (1) and (2). **Y* := Embed(Y,i*,x).**

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(k D^{\mu}_{1/5, 40\varepsilon/k}(f))$

This **i*** satisfies:

1. $E_{Y\sim\mu^k}[\ \Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}]\ ] \leq \varepsilon$
2. $E_{Y\sim\mu^k}[\Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 10\ \varepsilon/k$
3. $E_{Y\sim\mu^k}[\ E_X\ [q_{i^*}(Y^*)]\ ] \leq 3q/k$

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

This **i*** satisfies:

1. $E_{Y\sim\mu^k}[\ Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}]\ ] \leq \varepsilon$
2. $E_{Y\sim\mu^k}[Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 10\ \varepsilon/k$
3. $E_{Y\sim\mu^k}[\ E_X\ [q_{i^*}(Y^*)]\ ] \leq 3q/k$

Markov Inequality: there is **y*** such that

1. $Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}] \leq 4\varepsilon$
2. $Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 40\ \varepsilon/k$
3. $E_X\ [q_{i^*}(Y^*)] \leq 12q/k$

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

This **i*** satisfies:

1. $E_{Y\sim\mu^k}[\ Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}]\ ] \leq \varepsilon$
2. $E_{Y\sim\mu^k}[Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 10\ \varepsilon/k$
3. $E_{Y\sim\mu^k}[\ E_X\ [q_{i^*}(Y^*)]\ ] \leq 3q/k$

Markov Inequality: there is **y*** such that

1. $Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}] \leq 4\varepsilon$
2. $Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 40\ \varepsilon/k$
3. $E_X\ [q_{i^*}(Y^*)] \leq 12q/k$

```
Algorithm B(x) {
  z := EMBED(y*,i*,x)
  emulate A(z)
  abort if q_{i*}(z) > 120q/k
  abort if A(z)_{<i*} ≠ f^k(z)_{<i*}
}
```

Strong Direct Sum Theorem: $D^{\mu^k}_{0,\varepsilon}(f^k) = \Omega(kD^{\mu}_{1/5,40\varepsilon/k}(f))$

This **i*** satisfies:

1. $E_{Y\sim\mu^k}[\ Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}]\ ] \leq \varepsilon$
2. $E_{Y\sim\mu^k}[Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 10\ \varepsilon/k$
3. $E_{Y\sim\mu^k}[\ E_X\ [q_{i^*}(Y^*)]\ ] \leq 3q/k$

Markov Inequality: there is **y*** such that

1. $Pr_{x\sim\mu}[A(Y^*)_{<i^*} \neq f^k(Y^*)_{<i^*}] \leq 4\varepsilon$
2. $Pr_{x\sim\mu}[A(Y^*)_{i^*} \neq f^k(Y^*)_{i^*} \mid A(Y^*)_{<i^*} = f^k(Y^*)_{<i^*}] \leq 40\ \varepsilon/k$
3. $E_X\ [q_{i^*}(Y^*)] \leq 12q/k$

```
Algorithm B(x) {
  z := EMBED(y*,i*,x)
  emulate A(z)
  abort if qi*(z) > 120q/k
  abort if A(z)<i* ≠ fk(z)<i*
}
```

abort probability: $1/10 + 4\varepsilon < 1/5$

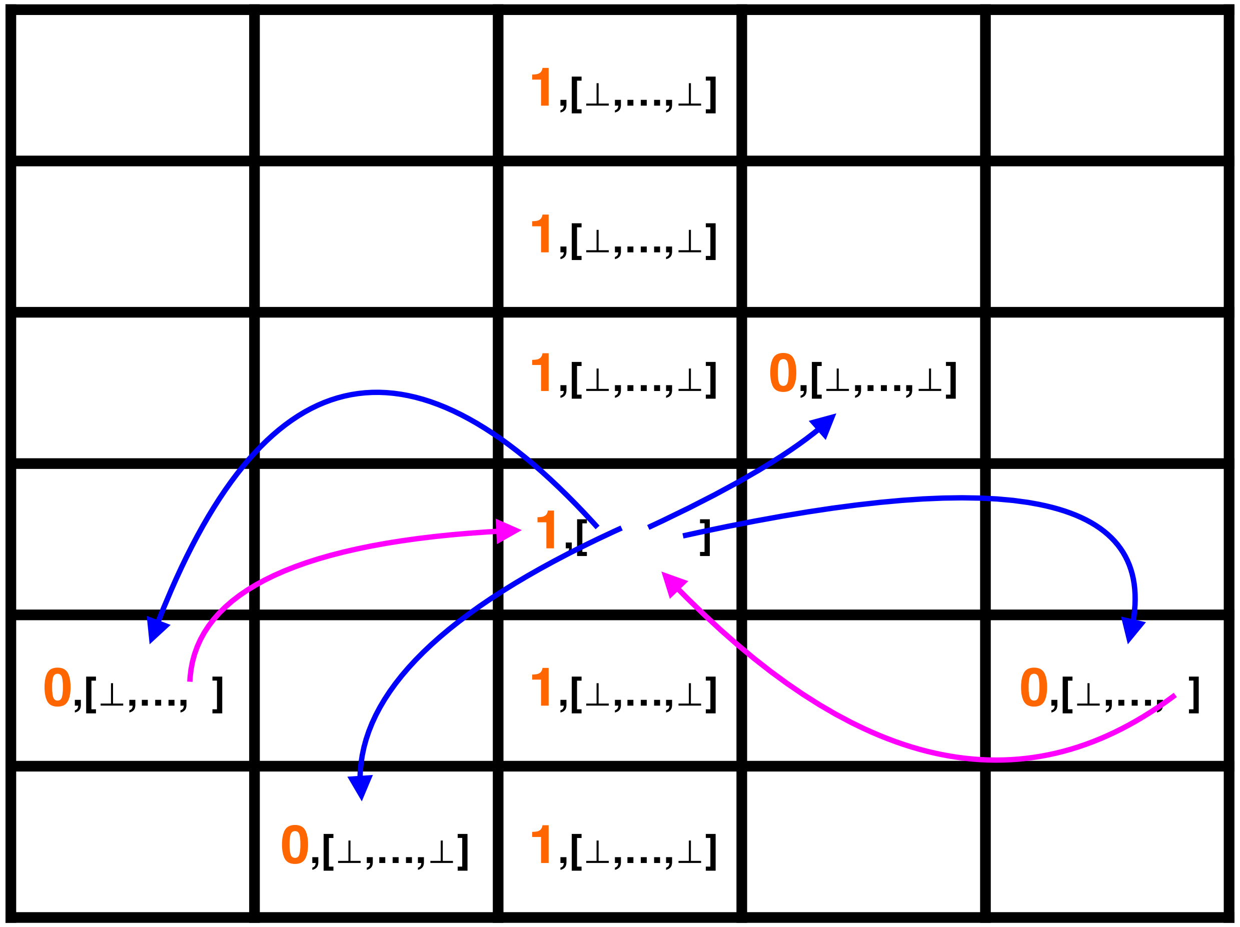
error probability: $40\varepsilon/k$

Outline

- Introduction
- Strong Direct Sum
- **Query Resistance**
- Separation Theorem
- Open Problems

Pointer Function

[GPW15, ABBLSS17,BB19]



PtrFcn: $\Sigma^{n*n} \rightarrow \{0,1\}$.

each cell **$z \in \Sigma$** has:

- value ***$b \in \{0,1\}$***
- **n** row ptrs ***$row_1(z), \ldots, row_n(z)$***
- back ptr ***back(z)***

PtrFcn(X) := 1 iff

- ∃ unique col j*: **$val(z_{i,j^*}) = 1$** for all i.
- ∃ *special cell* **z_{i^*,j^*}**. all ptrs **NULL** in col j* except for special cell
- special cell pts to **0**-value *linked cells* in each other col
- exactly half of *linked cells* point back to *special cell*

Query Resistant Codes

Definition: a *δ**N**-query resistant code* of **Σ** is a set of distribs **{G(x)}**

- For each **x ϵ Σ**, **G(x)** is a distribution on **$\{0,1\}^N$**
- **{ *support*(G(x)) : x ϵ Σ}** partition **$\{0,1\}^N$**
- For all **S** $\subseteq$ **[N]** with **|S|** $\leq$ δ**N**, distributions **$G(x)_{|S} = G(x')_{|S}$**
- *"decoding function"* **h(y) := x** iff **y ϵ support(G(x))**

Query Resistant Codes

Definition: a *δ**N**-query resistant code* of Σ is a set of distribs $\{G(x)\}$

- For each $x \in \Sigma$, $G(x)$ is a distribution on $\{0,1\}^N$
- $\{$ *support*$(G(x)) : x \in \Sigma\}$ partition $\{0,1\}^N$
- For all $S \subseteq [N]$ with $|S| \leq \delta N$, distributions $G(x)_{|S} = G(x')_{|S}$
- *"decoding function"* $h(y) := x$ iff $y \in \text{support}(G(x))$

Theorem: [Chor er al. 85] For any Σ, there is a $(N/3)$-query resistant code with $N = 12.5\log(|\Sigma|)$. Furthermore, conditional distributions $G(x)|_S$ are uniform.

Query Resistance

For $\mathbf{f : \Sigma^n \to \{0,1\}}$, define $\mathbf{F : \{0,1\}^{nN} \to \{0,1\}}$ as:

$$\mathbf{F(y_1,\ldots,y_n) := f(h(y_1),\ldots, h(y_n))}$$

Theorem: $\mathbf{R^{cell}_{\delta,\varepsilon}(f) \le (3/N)R_{\delta,\varepsilon}(F)}$

Query Resistance

For $f : \Sigma^n \rightarrow \{0,1\}$, define $F : \{0,1\}^{nN} \rightarrow \{0,1\}$ as:

$$F(y_1,\ldots,y_n) := f(h(y_1),\ldots, h(y_n))$$

Theorem: $R^{\text{cell}}_{\delta,\varepsilon}(f) \leq (3/N)R_{\delta,\varepsilon}(F)$

Proof: Let A be a (q, δ, ε)-algorithm for F.

```
Algorithm B(x₁,…,xₙ) {
  emulate A(G(x₁),…,G(xₙ))
  when A queries G(xᵢ) for k-th time:
    if k < N/3, sample G(xᵢ) cond. on prev. queries
    if k = N/3, sample xᵢ
    if k ≥ N/3, sample G(xᵢ) cond. on prev. history.
}
```

Open Problems

1. *Characterize* functions robust to **aborts**
2. **Strong Direct Sum** for Composed Functions
   (a) XOR Lemma
   (b) Strong Direct Sum for MAJ
3. How does $\mathbf{R}_{\delta,\epsilon}\mathbf{(f)}$ compare to other QC measures?

Thanks!

NOTE: Swarthmore has a tenure-track opening for fall 2020!