Local functions \( (\text{NC}^0) \)

- \( f : \{0,1\}^n \rightarrow \{0,1\} \) \( d \)-local:
  output depends on \( d \) input bits

- **Fact**: \( \text{Parity}(x) = 1 \iff \sum x_i = 1 \mod 2 \)
  is not \( n-1 \) local

- **Proof**: Flip any input bit \( \Rightarrow \) output flips ♦
Local generation of \((Y, \text{parity}(Y))\)

- **Theorem** [Babai ; Boppana Lagarias '87]

There is \(f : \{0,1\}^n \rightarrow \{0,1\}^{n+1}\), each bit 2-local

Distribution \(f(X) \equiv (Y, \text{parity}(Y)) \quad (X, Y \in \{0,1\}^n \text{ uniform})\)
Our message

• Complexity theory of distributions (as opposed to functions)

How hard is it to generate (a.k.a. sample) distribution $D$ given random bits?

E.g., $D = (Y, \text{parity}(Y))$, $D = W_k := \text{uniform n-bit with k 1's}$
This work:

- DNF
- $\text{AC}^0$
- $\text{NC}^0$
- local
Rest of talk

- Generating $W_k := \text{uniform n-bit with k 1's}$
  - $\text{Local } (\text{NC}^0)$
  - Decision tree

- Results for $(Y, b(Y))$

- Proof of local lower bound for $W_{n/2}$
Our results: local

- **Theorem**

  \[ f : \{0,1\}^n \rightarrow \{0,1\}^n \quad 0.1 \log n \text{ - local} \]

  \[ \Downarrow \]

  \[ f(X) \text{ at Statistical Distance} > 1 - n^{-\Omega(1)} \]

  from \( W_{n/2} = \text{uniform w/ weight } n/2 \)

- **Tight up to \( \Omega() \):** \( f(x) = x \)

- **Extends to** \( W_k, k \neq n/2, \text{ tight?} \)
Our results: succinct data structures

- **Problem:**
  Store \( k \)-subset \( S \subseteq \{1, 2, \ldots, n\} \) in \( u = \text{optimal} + r \) bits, answer “\( i \in S? \)” probing \( d \) bits.

- **Connection:**
  Solution \( \Rightarrow \) generate \( W_{|S|=k} \) d-local, Stat. Distance < 1 - 2^{-r}

- **Corollary:** Need \( r > \Omega(\log n) \) if \( d = 0.1 \log n \)
  First lower bound for \( |S| = n/2, n/4, \ldots \)
Decision tree model

- \( f : \{0,1\}^m \rightarrow \{0,1\}^n \) depth \( d \)
  
  Each output bit \( f_i \) is depth-\( d \) decision tree

- Depth \( d \leq 2^d \) local
Our results: decision trees

- **Theorem** \( f : \{0,1\}^* \rightarrow \{0,1\}^n \) depth < 0.1 \( \log n \)
  \[\Rightarrow\] Distance( \( f(X), W_{n/2} \) ) > 1/n

- Worse than \( 1 - n^{-\Omega(1)} \) lower bound for local

- **Fact** building on [Czumaj Kanarek Lorys Kutyłowski]
  \( \exists f : \text{depth } O(\log n) \) and Distance(\( f(X), W_{n/2} \) ) < 1/n
Rest of talk

- Generating $W_k := \text{uniform n-bit with k 1's}$
  - Local $(\text{NC}^0)$
  - Decision tree

- Results for $(Y, b(Y))$

- Proof of local lower bound for $W_{n/2}$
Our results for (Y, b(Y))

- **Theorem:** \( f : \{0,1\}^n \rightarrow \{0,1\}^{n+1} \)

  - 0.1 log \( n \)-local \( \Rightarrow \) Distance\( (f(X), (Y, Y \mod p > p/2)) > 0.49 \)
  
  - 0.1 log \( n \)-depth \( \Rightarrow \) Distance\( (f(X), (Y, \text{majority } Y)) > 1/n \)

- **Theorem** building on [Matias Vishkin, Hagerup]

  \( \exists f \text{ bounded-depth circuit } \text{AC}^0 : \)

  Distance\( (f(X), (Y, \text{majority } Y)) < 2^{-n} \)

- **Challenge:** explicit boolean b : AC\(^0\) can't generate (Y, b(Y))
Rest of talk

- Generating $W_k := \text{uniform n-bit with k 1's}$
  - Local ($\text{NC}^0$)
  - Decision tree

- Results for $(Y, b(Y))$

- Proof of local lower bound for $W_{n/2}$
Local lower bound

- **Theorem:** Let \( f : \{0,1\}^n \to \{0,1\}^n \) : \( d = 0.1 \log n \)-local.

  \[ \Rightarrow \exists T \subseteq \{0,1\}^n : \left| \Pr[f(x) \in T] - \Pr[W_{n/2} \in T] \right| > 1 - n^{-\Omega(1)} \]

- **Warm-up scenarios:**

  - \( f(x) = 000111 \) **Low-entropy**  \( T := \{000111\} \)

    \[ \left| \Pr[ f(x) \in T] - \Pr[W_{n/2} \in T] \right| = \left| 1 - |T| / \binom{n}{n/2} \right| \]

  - \( f(x) = x \) **“Anti-concentration”**  \( T := \{z : \sum_i z_i \neq n/2\} \)

    \[ \left| \Pr[ f(x) \in T] - \Pr[W_{n/2} \in T] \right| = \left| 1 - \Theta(1)/\sqrt{n} - 0 \right| \]
Proof

- Input $X = (X_1, X_2, \ldots, X_s, H)$

- Fix $H$. Output block $B_i$ depends only on bit $X_i$

- Many $B_i$ constant ($B_i(0,H) = B_i(1,H)$) ⇒ low-entropy

- Many $B_i$ depend on $X_i$ ($B_i(0,H) \neq B_i(1,H)$)

Idea: Independent ⇒ anti-concentration: sum $\neq n/2$ w.h.p.
If many $\text{weight}(B_i(0,H)) \neq \text{weight}(B_i(1,H))$, use

**Anti-concentration Lemma** [Littlewood Offord]

For $a_1, a_2, ..., a_s \neq 0$, any $c$, $\Pr_{X \in \{0,1\}^s} \left[ \sum_i a_i X_i = c \right] < 1/\sqrt{n}$

**Problem:** $B_i(0,H) = 100, B_i(1,H) = 010$

high entropy but no anti-concentration

**Fix:** want many blocks 000 : high entropy $\Rightarrow$ different weight
Conclusion

- Complexity of distributions = uncharted territory

- Lower bounds for $W_k := \text{uniform n-bit with k 1's}$
  - Local $\implies$ lower bound for storing sets efficiently
  - Decision tree

- Lower bounds for $(Y, b(Y))$, e.g. $(Y, \text{majority } Y)$
Generating $W_k := \text{uniform } n\text{-bit with } k \text{ 1's}$

- Local $(\mathsf{NC}^0)$
- Decision tree

Results for $(Y, b(Y))$

Proof of local lower bound for $W_{n/2}$
Our results: decision trees

• Theorem \( f : \{0,1\}^* \rightarrow \{0,1\}^n \) depth \( < 0.1 \log n \) \( \Rightarrow \) Distance( \( f(X) \), \( W_{n/2} \) ) \( > \) \( 1/n \)

• Proof: Is \( f(X) \) 4-wise independent?

  **YES:** [Paley Zygmund] \( \sum f(x)_i \) anti-concentrated, \( \neq n/2 \) w.h.p.

  **NO:** Let \( Q := \) biased 4 bits of \( f(X) \)

  Distance ( \( f(X) \mid_Q , W_{n/2} \mid_Q \approx \) uniform) \( > 2^{-4 \ (0.1 \log n)} \)

  by granularity of decision-tree probability
Test $T \subseteq \{0,1\}^n$: $\Pr[f(X_1,...,X_s,H) \in T] \approx 1$; $\Pr[W_{n/2} \in T] \approx 0$

$z \in T \iff$

$\exists H : \exists X_1,...,X_s$ w/ many blocks $B_i$ fixed: $f(X_1,...,X_s,H) = z$

OR

Few blocks $z|_{B_i}$ are 000

OR

$\Sigma_i z_i \neq n/2$
Rest of this talk

- Connection with succinct data structures
- Lower bound for locally generating $W_{n/2} = \text{n-bit with n/2 1's}$
- Decision tree model
- Bounded-depth circuit model
Tool for lower bound proof

- Central limit theorem:
  \[ x_1, x_2, ..., x_n \text{ independent} \implies \sum x_i \approx \text{normal} \]

- Bounded-independence central limit theorem
  [Diakonikolas Gopalan Jaiswal Servedio V.]
  \[ x_1, x_2, ..., x_n \text{ k-wise independent} \implies \sum x_i \approx \text{normal} \]

- Note: For next result, Paley–Zygmund inequality enough
Proof

- **Theorem [V.]** \( f: \{0,1\}^* \rightarrow \{0,1\}^n \): each bit depth \(< 0.1 \log n\)
  \[\text{Distance}(f(X), W_{n/2}) > n^{-\Omega(1)}\]

- **Proof:** Is output distribution \( f(X) \) \((k = 10)\)-wise independent?

  **NO** \(\Rightarrow W_{n/2} \approx k\)-wise independent
  \[\text{Distance(those k bits, uniform on } \{0,1\}^k) > 2^{-k(0.1 \log n)}\]
  (granularity of decision tree probability)

  **YES** \(\Rightarrow\) by prev. theorem \(\sum f(X)_i \approx \text{normal}\)
  so often \(\sum f(X)_i \neq n/2\)
Rest of this talk

- Connection with succinct data structures
- Lower bound for locally generating $W_{n/2} = n$-bit with $n/2$ 1's
- Decision tree model
- Bounded-depth circuit model
Lower bound for codes

- **Code** $C \subseteq \{0,1\}^n$ of size $|C| = 2^k = \Omega(n)$
  
  $x \neq y \in C \implies x, y$ far : hamming distance $\Omega(n)$

- **Theorem** [Lovett V.] $f : \{0,1\}^* \rightarrow \{0,1\}^n$, $f \in AC^0$
  
  Distance($f(X)$, uniform over $C$) $> 1 - n^{-\Omega(1)}$

- Consequences for data structures for codewords, complexity of pseudorand. generators against $AC^0$ [Nisan]
Warm-up

- Fact: $f : \{0,1\}^k \rightarrow \{0,1\}^n$, $f \in AC^0$
  $f$ cannot compute encoding function of $C$
  mapping message $m \in \{0,1\}^k$ to codeword

- Proof:

- [Linial Mansour Nisan, Boppana] low sensitivity of $AC^0$:
  $m, m'$ random at hamming distance 1
  $\Rightarrow f(m), f(m')$ close in hamming distance.

- But $f(m) \neq f(m') \in C \Rightarrow$ far in hamming distance
Lower bound for codes

• **Theorem [Lovett V.]** \( f : \{0,1\}^L \gg k \rightarrow \{0,1\}^n \), \( f \in \text{AC}^0 \)
  
  Distance\((f(X), \text{uniform over } C) > 1 - n^{-\Omega(1)} \)

**Problem:** \( f \) needs not compute encoding function.
Input length \( \gg \) message length

• **Idea:** Input \( \{0,1\}^L \) to \( f \) partitioned in \( |C| \) sets

• **Isoperimetric inequality [Harper, Hart]:**
  Random \( m, m' \) at distance 1 often in \( \neq \) sets \( \Rightarrow \) low sensitivity
Lower bound for codes

- **Theorem [Lovett V.]** \( f : \{0,1\}^L \gg k \rightarrow \{0,1\}^n \), \( f \in AC^0 \)
  \[ \text{Distance}(f(X), \text{uniform over } C) > 1 - n^{-\Omega(1)} \]

- **Note:** to get
  Need isoperimetric inequality for \( m, m' \) at distance \( \gg 1 \)

**Fact**[thanks to Samorodnitsky] \( \forall A \subseteq \{0,1\}^L \) of density \( \alpha \)
random \( m, m' \) obtained flipping bits w/ probability \( p \):
\[ \alpha^2 \leq \Pr[\text{both } m \in A \text{ and } m' \in A] \leq \alpha^{1/(1-p)} \]
• $\sum \cap \neg \cup \mathfrak{a} \beta \gamma \delta \rightarrow$
• $\neq \Theta \Omega \theta$

• Recall: edit style changes ALL settings.
• Click on “line” for just the one you highlight
More connections

- More uses of generating $W_k := \text{uniform n-bit string with k 1's}$
- McEliece cryptosystem
- Switching networks, …
Previous results

- Store $S \subseteq \{1, 2, \ldots, n\}$, $|S| = k$, in bits, answer “$i \in S$?”

- [Minsky Papert '69] Average-case study

- [Buhrman Miltersen Radhakrishnan Venkatesh; Pagh '00]
  
  Space $O(\text{optimal})$, probe $O(1)$ when $k = \Theta(n)$

  Lower bounds for $k < n^{1-\varepsilon}$

- [...] Pagh, Pătrașcu] space = optimal + $o(n)$, probe $O(\log n)$

- [V. '09] lower bounds for $k = \Omega(n)$, except $k = n / 2^a$
Succinct data structures for sets

- Store $S \subseteq \{1, 2, \ldots, n\}$ of size $|S| = k$

- In $u$ bits $b_1, \ldots, b_u \in \{0,1\}$

- Want:
  - Small space $u$ (optimal $= \lceil \lg_2 \binom{n}{k} \rceil$)
  - Answer “$i \in S$?” by probing few bits (optimal $= 1$)

- In combinatorics: Nešetřil Pultr, …, Körner Monti
Previous results

- Store $S \subseteq \{1, 2, \ldots, n\}$, $|S| = k$, in bits, answer “$i \in S$?”

- [Minsky Papert '69, Buhrman Miltersen Radhakrishnan Venkatesh; Pagh; ...; Pătraşcu; V. '09]

- Surprising upper bounds
  space = optimal + o(n), probe $O(\log n)$

- No lower bounds for $k = n / 2^a$
Rest of this talk

- Local \((NC^0)\)
  
  Lower bound for \(W_{n/2} = n\)-bit with \(n/2\) 1's
  
  Succinct data structures

- Decision tree
  
  Lower bound for \(W_{n/2}\)

- Bounded-depth circuit \((AC^0)\)

- Proof of local lower bound
Bounded-depth circuits \((\text{AC}^0)\)

- \(O(\log n)\)-local \(\subseteq\) depth \(O(\log n)\) \(\subseteq\) \(\text{AC}^0\)

- **Theorem** [Matias Vishkin, Hagerup, this work]
  Can generate \(W_k\), exp. small error

- **Theorem** [Lovett V.] **Cannot** generate error-correcting code

- **Challenge:** \(\exists\) explicit boolean \(f\) : cannot generate \((Y, f(Y))\) ?
Our results: pseudorandomness for $\text{AC}^0$

- Pseudorandom distribution against circuit of depth $d$
  (want: reduce randomness w/ minimum overhead)

- **Direct implementation** of Nisan's generator: depth $\geq d$
  circuit + generator $\rightarrow$ depth $2d$

- **Generator in depth 2** circuit + generator $\rightarrow$ depth $d+1$
  [Braverman] + [Guruswami Umans Vadhan]