Quantum Computing,

Locally Decodable Codes,

and

Private Information Retrieval

Iordanis Kerenidis (UC Berkeley)

Ronald de Wolf (CWI Amsterdam)

Error-Correcting Codes

- Encoding $C : \{0, 1\}^n \to \{0, 1\}^m$, $m \geq n$

- Even if $C(x)$ is corrupted in δm positions, we can still recover the whole x

- We can achieve this with $m = O(n)$, linear-time encoding and decoding.
  $O(1)$ time per bit!

- Disadvantage: if you only want one bit x_i, you still need to decode the whole $C(x)$

Locally Decodable Codes

- Recover x_i with high probability, looking only at a few positions in the codeword

- $C : \{0,1\}^n \to \{0,1\}^m$ is a (q, δ, ε)-*locally decodable code* (LDC) if there exists a randomized decoder A such that for every $y \in \{0,1\}^m$ and $i \in [n]$

  1. $A^y(i)$ makes $\leq q$ queries to bits of y (non-adaptively)

  2. $d(y, C(x)) \leq \delta m \Rightarrow \Pr[A^y(i) = x_i] \geq 1/2 + \varepsilon$

- LQDCs: classical code, quantum queries

Example: Hadamard Code

- Define $C(x)_j = j \cdot x \bmod 2$ for all $j \in \{0,1\}^n$, so $m = 2^n$
- Decode: pick random $j \in \{0,1\}^n$, query j and $j \oplus e_i$, output $y_j \oplus y_{j \oplus e_i}$
- Works perfectly if $y = C(x)$ (no noise)
- δ-corruption hits $C(x)_j$ or $C(x)_{j \oplus e_i}$ with probability $\leq 2\delta$, so

$$\Pr[A^y(i) = x_i] \geq 1 - 2\delta$$

What Was Known About LDCs

Main question: tradeoff between q and m

- Upper bounds:

  $q = m \Rightarrow m \leq O(n)$ (standard ECC)
  $q = (\log n)^2 \Rightarrow m \leq poly(n)$ (Babai et al)
  constant $q \Rightarrow m \leq 2^{n^{c(q)}}$ (from PIR)

- Lower bounds:

  Katz-Trevisan 00:
  $q = 1 \Rightarrow$ LDCs don't exist
  $q > 1 \Rightarrow m \geq n^{1+\frac{1}{q-1}}$
  GKST 02:
  $q = 2$, linear $C \Rightarrow m \geq 2^{cn}, c = \delta\varepsilon/8$

- Our result:

  $q = 2 \Rightarrow m \geq 2^{c'n}$ also for non-linear LDCs

Our Proof Uses Quantum!

- Step 1:

  2-query LDCs can be decoded
  with 1 quantum query:

  $(2, \delta, \varepsilon)$-LDC is $(1, \delta, 4\varepsilon/7)$-LQDC

  (example: Hadamard code)

- Step 2:

  $(1, \delta, \varepsilon)$-LQDC needs length $m \geq 2^{c'n}$,

  because it implies a random access code

Step 1: From 2-LDC to 1-LQDC

Compute Boolean function $f(a_1, a_2)$ with 1 quantum query and success probability exactly 11/14:

1. Query $|\phi\rangle = |0\rangle + (-1)^{a_1}|1\rangle + (-1)^{a_2}|2\rangle$

2. Measure in 4-element basis $|\psi_{b_1 b_2}\rangle =$ $|0\rangle + (-1)^{b_1}|1\rangle + (-1)^{b_2}|2\rangle + (-1)^{b_1+b_2}|3\rangle$

3. $\Pr[b_1 b_2 = a_1 a_2] = |\langle\phi|\psi_{a_1 a_2}\rangle|^2 = 3/4$

4. $b_1 b_2$ + truth table of $f \Rightarrow$ output

For classical 2-query decoder with success probability $p = 1/2 + \varepsilon$, one quantum query gives

$$\frac{11}{14}p + \frac{3}{14}(1-p) = \frac{1}{2} + \frac{4\varepsilon}{7}$$

Step 2: Lower Bound for 1-LQDC

- Quantum decoder predicts x_i by doing POVM on query state $\sum_{j=1}^{m}(-1)^{C(x)_j}\alpha_j|j\rangle$

- This can tolerate up to δm phase-errors

- Small amplitudes $A_i = \{j : \alpha_j \leq 1/\sqrt{\delta m}\}$ misses at most δm indices

- Given $|A_i(x)\rangle = \sum_{j \in A_i}(-1)^{C(x)_j}\alpha_j|j\rangle$,
  we can predict x_i with good bias $\approx \varepsilon$

Step 2: get $|A_i(x)\rangle$ from uniform state

- Predict x_i from $|U(x)\rangle = \sum_{j=1}^{m} (-1)^{C(x)_j} |j\rangle$:

  1. Measure $|U(x)\rangle$ with POVM $M_i^* M_i$, $I - M_i^* M_i$, where $M_i = \sqrt{\delta m} \sum_{j \in A_i} \alpha_j |j\rangle\langle j|$

  2. With prob $\approx \delta$: $M_i : |U(x)\rangle \mapsto |A_i(x)\rangle$, then we can predict x_i with bias $\approx \varepsilon$

     With prob $\approx 1 - \delta$: output fair coin flip

  3. This gives x_i with prob $p \approx 1/2 + \delta\varepsilon$

- $|U(x)\rangle$ is a random access code for x!

$$\underbrace{\log m}_{\#\text{qubits of } U(x)} \geq \underbrace{(1 - H(p))n}_{\text{RAC bound (Nayak 99)}}$$

LQDCs are shorter than LDCs

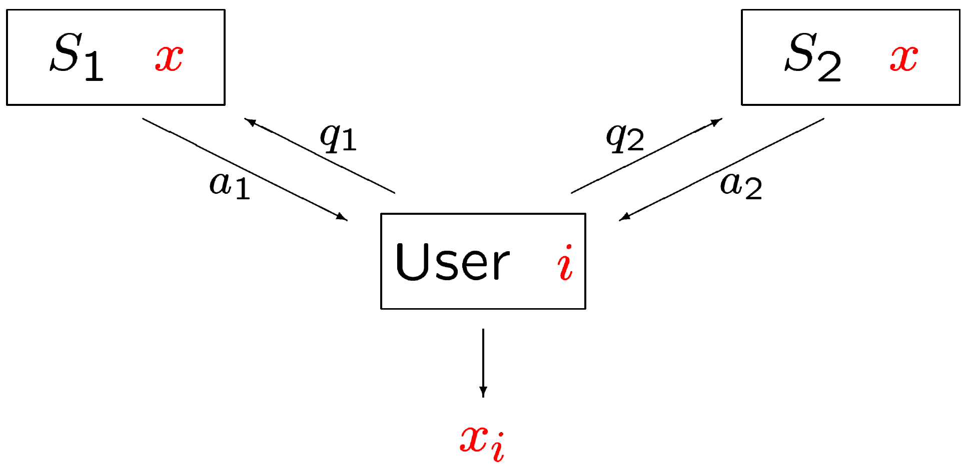
- Best known $2q$-query LDCs (BIKR 02) output the XOR of the $2q$ bits

- Can do this with q quantum queries!

Queries	Length of LDC	Length of LQDC
$q = 1$	don't exist	$2^{\Theta(n)}$
$q = 2$	$2^{\Theta(n)}$	$2^{n^{3/10}}$
$q = 3$	$2^{n^{1/2}}$	$2^{n^{1/7}}$
$q = 4$	$2^{n^{3/10}}$	$2^{n^{1/11}}$

Private Information Retrieval

- User retrieves x_i with probability $1/2 + \varepsilon$ from n-bit database x that is replicated over k non-communicating servers



- Privacy: server learns nothing about i

- How much communication is needed?
  - 1-server PIR needs $\Omega(n)$ bits
  - 2-server PIR with $O(n^{1/3})$ bits (CGKS)

Lower Bound for Classical Binary PIR

- Binary PIR: servers send back only 1 bit
- Can reduce 2 binary classical servers to 1 quantum server (treat servers as queries)
- $\Omega(n)$ lower bound for 1-server quantum PIR
  $\Rightarrow$
  $\Omega(n)$ lower bound for 2-server binary PIR
- Previously known only for *linear* PIR (GKST)
- Recent classical proof if $\varepsilon = 1/2$ (BFG)

Upper Bound for Quantum PIR

- Best known $2k$-server binary PIRs (BIKR 02) output XOR of the $2k$ bits
- Can do this with k quantum servers
- Better than best known k-server PIRs!

Servers	PIR complexity	QPIR complexity
$k = 1$	n	n
$k = 2$	$n^{1/3}$	$n^{3/10}$
$k = 3$	$n^{1/5.25}$	$n^{1/7}$
$k = 4$	$n^{1/7.87}$	$n^{1/11}$

Summary

- Locally decodable codes:
  - Exponential lower bound for 2-query LDCs via a quantum proof
  - q-query LQDCs are shorter than LDCs
- Private information retrieval:
  - $\Omega(n)$ lower bound for 2-server binary PIR
  - Upper bound $O(n^{3/10})$ for 2-server QPIR