Tight Quantum Lower Bound for k-Distinctness

Alexander Belov
University of Latvia

Joint Estonian-Latvian Theory Days
Tartu, April 26th, 2026

Problem

k-Distinctness

- k-Distinctness: Given a string

$$x = (x_1, \ldots, x_n)$$

  find k elements that are equal.
- Case $k = 2$ is known as Element Distinctness.
- Inspired a number of techniques in upper and lower bounds.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf. $\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.

- Using quantum amplitude amplification.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.
2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.

- Using the polynomial method.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.
2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.
2003 Ambainis.
$\mathcal{O}(n^{k/(k+1)})$ upper bound (time and query).

- Using brand-new quantum walk on the Johnson graph.
- Tight $\Theta(n^{2/3})$ for Element Distinctness.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.

2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.

2003 Ambainis.
$\mathcal{O}(n^{k/(k+1)})$ upper bound (time and query).

2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.

- Using a modification of Leaning Graphs.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.
2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.
2003 Ambainis.
$\mathcal{O}(n^{k/(k+1)})$ upper bound (time and query).
2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.
2013 Belovs, Childs, Jeffery, Kothari, Magniez.
Time-efficient $\mathcal{O}(n^{5/7})$ algorithm for $k = 3$.

- Using brand-new electric quantum walks.
- Using brand-new nested quantum walks.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.

2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.

2003 Ambainis.
$\mathcal{O}(n^{k/(k+1)})$ upper bound (time and query).

2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.

2013 Belovs, Childs, Jeffery, Kothari, Magniez.
Time-efficient $\mathcal{O}(n^{5/7})$ algorithm for $k = 3$.

2017-20 Bun, Kothari, Thaler; Mande, Thaler, Zhu
$\Omega(n^{\frac{3}{4}-\frac{1}{2k}})$ and $\Omega(n^{\frac{3}{4}-\frac{1}{4k}})$ lower bounds.

- Using dual polynomials.

History

2000 Buhrman, Dürr, Heiligman, Høyer, Magniez, Santha, de Wolf.
$\mathcal{O}(n^{3/4})$ query and time algorithm for Element Distinctness.
2002 Aaronson, Shi.
$\Omega(n^{2/3})$ lower bound for Element Distinctness.
2003 Ambainis.
$\mathcal{O}(n^{k/(k+1)})$ upper bound (time and query).
2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.
2013 Belovs, Childs, Jeffery, Kothari, Magniez.
Time-efficient $\mathcal{O}(n^{5/7})$ algorithm for $k = 3$.
2017-20 Bun, Kothari, Thaler; Mande, Thaler, Zhu
$\Omega(n^{\frac{3}{4}-\frac{1}{2k}})$ and $\Omega(n^{\frac{3}{4}-\frac{1}{4k}})$ lower bounds.
2022 Jeffery, Zur.
Time-efficient $\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ algorithm.

- Using brand-new multidimensional quantum walks.

This work

2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.

2020 Mande, Thaler, Zhu.
$\Omega\left(n^{\frac{3}{4}-\frac{1}{4k}}\right)$ lower bound.

2022 Jeffery, Zur.
Time-efficient $\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ algorithm.

This work

2012 Belovs.
$\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ query upper bound.
2020 Mande, Thaler, Zhu.
$\Omega(n^{\frac{3}{4}-\frac{1}{4k}})$ lower bound.
2022 Jeffery, Zur.
Time-efficient $\mathcal{O}\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$ algorithm.

This work: Tight

$$\Omega\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\right)$$

query lower bound for any k assuming the alphabet size $q \gg \Omega(n^2)$.

- Using brand-new lower bound framework.

Techniques

- Aaronson, Shi, Bun, Kothari, Thaler, Mande, Zhu all used the polynomial method.

Techniques

- Aaronson, Shi, Bun, Kothari, Thaler, Mande, Zhu all used the polynomial method.

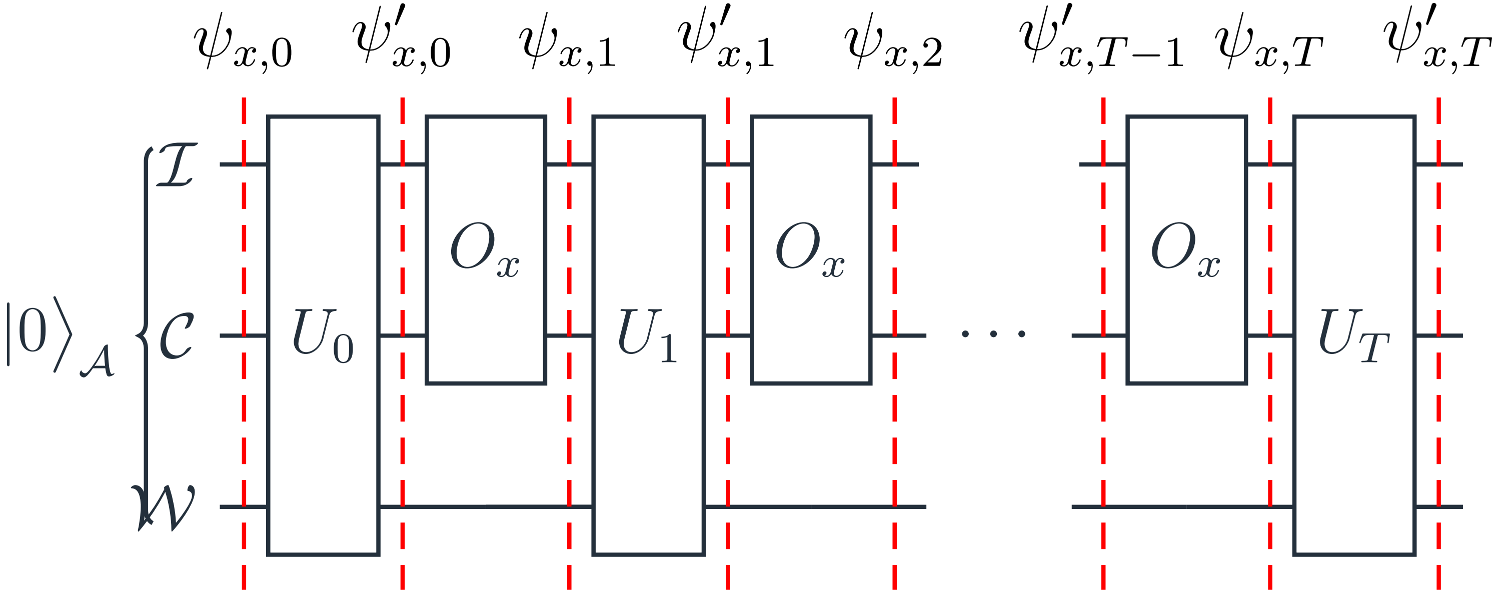
- Important method is Compressed Oracle Technique due to Zhandry, 2017.
  - ☐ Is rather intuitive.
  - ☐ Restricted to uniform probability on the input strings.
  - ☐ Actually, Liu and Zhandry characterise complexity of finding k-collision in a uniformly random string.

Techniques

- Aaronson, Shi, Bun, Kothari, Thaler, Mande, Zhu all used the polynomial method.
- Important method is Compressed Oracle Technique due to Zhandry, 2017.
  - Is rather intuitive.
  - Restricted to uniform probability on the input strings.
  - Actually, Liu and Zhandry characterise complexity of finding k-collision in a uniformly random string.
- We use a different technique that
  - is a generalisation of the polynomial method,
  - inspired by Zhandry's ideas,
  - works for non-uniform distributions.

Pillars

Quantum Query Algorithm

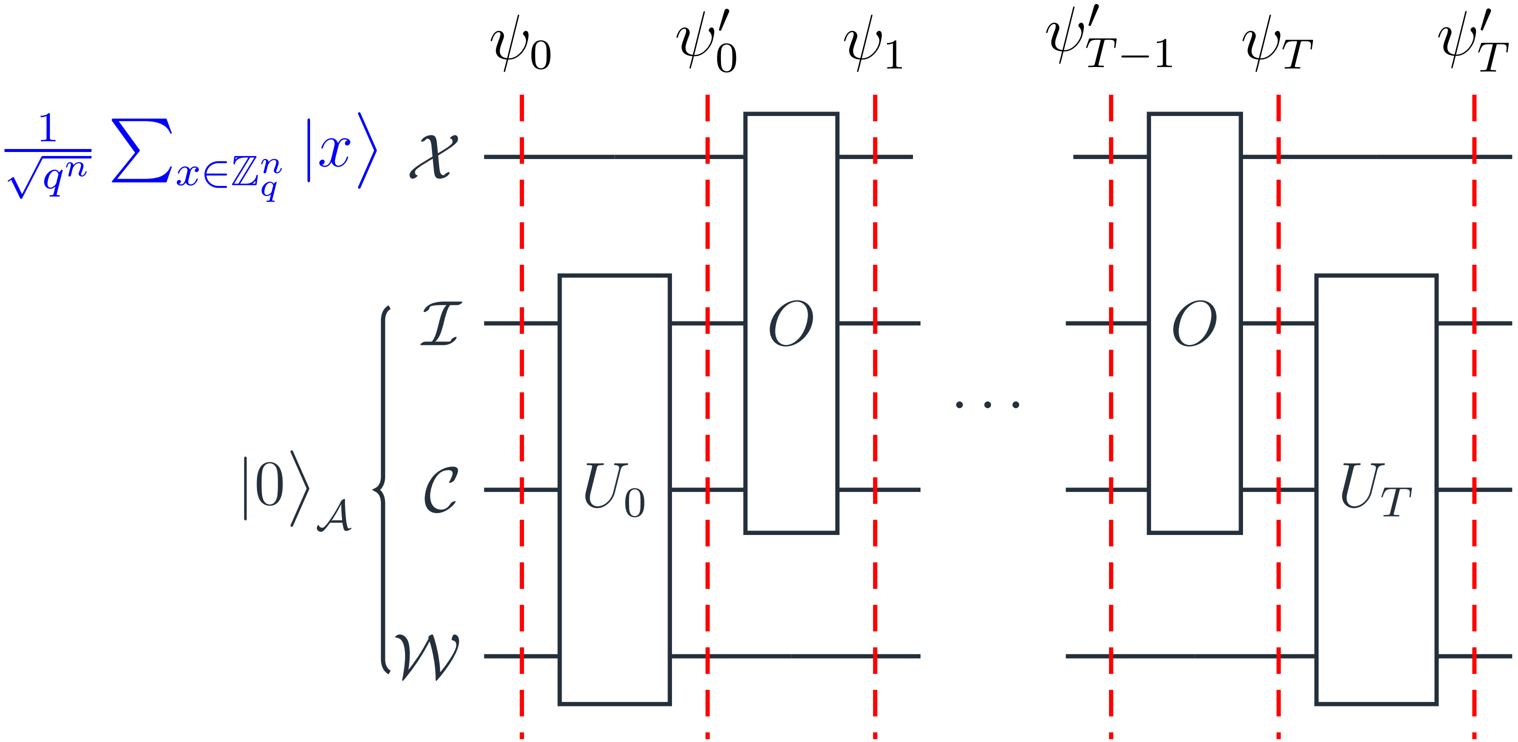


- Alternates between unitaries U_t and queries O_x in the phase:

$$O_x : \quad |i\rangle_{\mathcal{I}} |c\rangle_{\mathcal{C}} \mapsto \omega_q^{cx_i} |i\rangle_{\mathcal{I}} |c\rangle_{\mathcal{C}}.$$

- States of the algorithm: $\psi_{x,t}$ and $\psi'_{x,t}$.

Algorithm Combined with Input



- **Uniform States of the Algorithm ψ_t and ψ'_t:**

$$\psi_t = \frac{1}{\sqrt{q^n}}\sum_{x\in\mathbb{Z}_q^n}|x\rangle_{\mathcal{X}}|\psi_{x,t}\rangle_{\mathcal{A}}.$$

Phase Kickback

Proposition

The input oracle O in the Fourier basis acts as

$$O\colon |\widehat{\sigma}\rangle_{\mathcal{X}} |i\rangle_{\mathcal{I}} |c\rangle_{\mathcal{C}} \mapsto |\widehat{\sigma + \{i \mapsto c\}}\rangle_{\mathcal{X}} |i\rangle_{\mathcal{I}} |c\rangle_{\mathcal{C}}.$$

where $\{i \mapsto c\} \in \mathbb{Z}_q^n$ is a function

$$j \mapsto \begin{cases} c, & \text{if } j = i; \\ 0, & \text{otherwise;} \end{cases}$$

and addition of function is performed element-wise.

Phase Kickback

Proposition

The input oracle O in the Fourier basis acts as

$$O\colon |\widehat{\sigma}\rangle_{\mathcal{X}}|i\rangle_{\mathcal{I}}|c\rangle_{\mathcal{C}} \mapsto |\widehat{\sigma + \{i \mapsto c\}}\rangle_{\mathcal{X}}|i\rangle_{\mathcal{I}}|c\rangle_{\mathcal{C}}.$$

- Oracle acts in $\mathcal{X}$, not in $\mathcal{A}$, which we have much more control over.
- Gives important piece of intuition: the state $|\widehat{\sigma}\rangle_{\mathcal{X}}$
  - has knowledge about the variables in the support of σ;
  - has no knowledge about the variable outside of it.

Phase Kickback

Proposition

The input oracle O in the Fourier basis acts as

$$O\colon |\hat{\sigma}\rangle_{\mathcal{X}}|i\rangle_{\mathcal{I}}|c\rangle_{\mathcal{C}} \mapsto |\widehat{\sigma + \{i \mapsto c\}}\rangle_{\mathcal{X}}|i\rangle_{\mathcal{I}}|c\rangle_{\mathcal{C}}.$$

Corollary

After t queries, the uniform state of the algorithm

$$\psi_t, \psi'_t \in \mathcal{X}_{\leq t} \otimes \mathcal{A},$$

where $\mathcal{X}_{\leq t}$ is spanned by $|\hat{\sigma}\rangle_{\mathcal{X}}$ with support size $\leq t$.

- The algorithm knows values of some $\leq t$ variables.

Non-Uniform Distribution

Transfer Operator

- So far, we used uniform probability distribution on the input.
- What if we want another probability distribution $p = (p_x)$?

Transfer Operator

- So far, we used uniform probability distribution on the input.
- What if we want another probability distribution $p = (p_x)$?

- Let $\mathcal{Y}$ be another register isomorphic to $\mathcal{X}$.
- Define transfer operator

$$\Upsilon\colon \mathcal{X} \to \mathcal{Y}\colon \quad \frac{1}{\sqrt{q^n}}|x\rangle_{\mathcal{X}} \mapsto \sqrt{p_x}|x\rangle_{\mathcal{Y}}.$$

Transfer Operator

- So far, we used uniform probability distribution on the input.
- What if we want another probability distribution $p = (p_x)$?

- Let $\mathcal{Y}$ be another register isomorphic to $\mathcal{X}$.
- Define transfer operator

$$\Upsilon\colon \mathcal{X} \to \mathcal{Y}\colon \qquad \frac{1}{\sqrt{q^n}} |x\rangle_{\mathcal{X}} \mapsto \sqrt{p_x} |x\rangle_{\mathcal{Y}}.$$

- It is obvious that

$$\psi_t = \frac{1}{\sqrt{q^n}} \sum_{x \in \mathbb{Z}_q^n} |x\rangle_{\mathcal{X}} |\psi_{x,t}\rangle_{\mathcal{A}} \xmapsto{\Upsilon} \Upsilon\psi_t = \sum_{x \in \mathbb{Z}_q^n} \sqrt{p_x} |x\rangle_{\mathcal{Y}} |\psi_{x,t}\rangle_{\mathcal{A}}.$$

Anti-Concentration

- For $0 < \gamma < 1$:
  - ☐ vector $\phi \in \mathcal{Y}$ is γ-anti-concentrated if $\|M_\rho \phi\| \leq \gamma \|\phi\|$ for all $\rho \in R$.
  - ☐ subspace $\mathcal{H} \subseteq \mathcal{Y}$ is γ-anti-concentrated if all $\phi \in \mathcal{H}$ are.
  - ☐ vector $\psi \in \mathcal{Y} \otimes \mathcal{A}$ is strongly γ-anti-concentrated if $\psi \in \mathcal{H} \otimes \mathcal{A}$ with $\mathcal{H}$ being γ-anti-concentrated.

Anti-Concentration

- For $0 < \gamma < 1$:
  - ☐ vector $\phi \in \mathcal{Y}$ is γ-anti-concentrated if $\|M_\rho \phi\| \leq \gamma \|\phi\|$ for all $\rho \in R$.
  - ☐ subspace $\mathcal{H} \subseteq \mathcal{Y}$ is γ-anti-concentrated if all $\phi \in \mathcal{H}$ are.
  - ☐ vector $\psi \in \mathcal{Y} \otimes \mathcal{A}$ is strongly γ-anti-concentrated if $\psi \in \mathcal{H} \otimes \mathcal{A}$ with $\mathcal{H}$ being γ-anti-concentrated.

Proposition

If the final state $\Upsilon \psi'_T$ of the algorithm is strongly γ-anti-concentrated, then the average (w.r.t. p) success probability of the algorithm is at most γ^2.

Polynomial Method

- The polynomial method proves lower bounds on quantum query complexity of a (partial) Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ by showing that f cannot be approximated by a low degree polynomial.

Polynomial Method

- The polynomial method proves lower bounds on quantum query complexity of a (partial) Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ by showing that f cannot be approximated by a low degree polynomial.

- This is a special case of the above construction.
- The subspace $\Upsilon \mathcal{X}_{\le t} \subseteq \mathcal{Y}$ is $\frac{1}{\sqrt{2}}$-anti-concentrated for a carefully constructed p coming from a dual polynomial.

Polynomial Method

- The polynomial method proves lower bounds on quantum query complexity of a (partial) Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ by showing that f cannot be approximated by a low degree polynomial.

- This is a special case of the above construction.
- The subspace $\Upsilon \mathcal{X}_{\le t} \subseteq \mathcal{Y}$ is $\frac{1}{\sqrt{2}}$-anti-concentrated for a carefully constructed p coming from a dual polynomial.

- Such p are very unnatural and contrived.
- We would like to use a natural p.

Knowledge

- For every input string (rather a subset μ of inputs), we define knowledge system L_μ^+ that consists of subsets of $[n]$ that have crucial knowledge about (every input with type) μ.

Knowledge

- For every input string (rather a subset μ of inputs), we define knowledge system L_μ^+ that consists of subsets of $[n]$ that have crucial knowledge about (every input with type) μ.

- Recall, for k-distinctness, types are given by partitions of $[n]$.
- $i, j \in [n]$ are in the same block of μ iff $y_i = y_j$.

- $S \in L_\mu^+$ iff S has intersection of size $\geq k$ with a block of μ.
- This means that S contains k equal elements for every input of type μ.

Knowledge Operator

- For every input string (rather a subset μ of inputs), we define knowledge system L_μ^+ that consists of subsets of $[n]$ that have crucial knowledge about (every input with type) μ.

- This allows us to define knowledge operator

$$\Upsilon_\mu^+ : \mathcal{X} \to \mathcal{Y}_\mu : \quad |\widehat{\sigma}\rangle_{\mathcal{X}} \longmapsto \begin{cases} \Upsilon_\mu |\widehat{\sigma}\rangle_{\mathcal{X}}, & \text{if } \operatorname{supp}(\sigma) \in L_\mu^+; \\ 0, & \text{otherwise.} \end{cases}$$

- In other words
  - ☐ we use Fourier decomposition in the uniform space $\mathcal{X}$ to define knowledge, and
  - ☐ we transfer it to $\mathcal{Y}_\mu$ using the transfer operator.

Framework

- This gives us a decomposition

$$\Upsilon\psi_t = \Upsilon^+\psi_t + \Upsilon^-\psi_t,$$

where $\Upsilon^- : \mathcal{X} \to \mathcal{Y}$ is the complement of Υ^+.

Framework

If we can show that

- (Small knowledge) $\|\Upsilon^+\psi'_T\| \leq \delta$ and
- (Anti-concentration) $\Upsilon^-\mathcal{X}_{\leq T}$ is γ-anti-concentrated,

then the average success probability of the algorithm is at most $(\gamma+\delta)^2$.

Search For Equal Elements

Formulation

Is given by a set M of partitions μ of $[n]$.

- Recall that partitions $\mu \in M$ of $[n]$ are related to inputs $y \in \mathbb{Z}_q^n$ by i, j share a block of μ iff $y_i = y_j$.

- We relax this condition and consider $Y_\mu = \mathbb{Z}_q^\mu$.
- If $q \gg n^2$, there is not much difference between the strict and the relaxed versions.

Advantages:

- We can introduce Fourier basis $|\mu, \hat{\tau}\rangle_{\mathcal{Y}}$ with $\tau \in \mathbb{Z}_q^\mu$ in $\mathcal{Y}$.
- The transfer operator has a very nice form:

$$\Upsilon_\mu \colon |\hat{\sigma}\rangle_{\mathcal{X}} \mapsto |\mu, \hat{\tau}\rangle_{\mathcal{Y}} \quad \text{with} \quad \tau(B) = \sum_{i \in B} \sigma(i) \quad \text{for all } B \in \mu.$$

Anti-Concentration

Theorem

If the set of partitions M is

- symmetric under permutations and
- has $\geq cn$ singletons in each μ,

then $\Upsilon^{-}\mathcal{X}_{\leq t}$ is $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$-anti-concentrated for each $t \leq cn/2$.

Thank you!

Thank you!

Supported by the Latvian Quantum Initiative under European Union Recovery and Resilience Facility project no. 2.3.1.1.i.0/1/22/I/CFLA/001.