# Quantum Query Complexity of Boolean Functions with Small On-Sets

Seiichiro Tani
NTT/JST ERATO-SORST.

Joint work with

Andris Ambainis
Univ. of Latvia

Kazuo Iwama
Kyoto Univ.

Masaki Nakanishi
NAIST.

Harumichi Nishimura
Osaka Pref. Univ

Rudy Raymond
IBM

Shigeru Yamashita
NAIST.

# Motivation

- Want to test some properties of huge data X,
  Or, compute some function f(X).
  - e.g. WWW log analysis, Experimental data analysis….

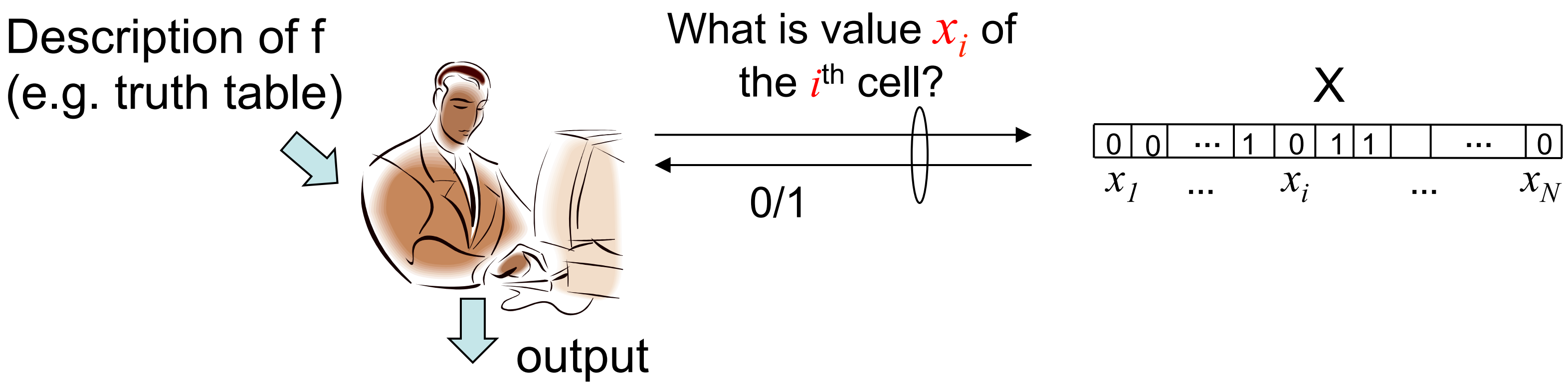
X

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | … | 1 | 0 | 1 | 1 | 0 | … | 0 |
| 1 | 2 | … | 101 | 102 | 103 | 104 | 105 | … | **N** |

- Reading all memory cells of X costs too much.
- Can we save the number of accessing X when computing certain functions f(X) ?

# Oracle Computation Model

Can know the value of one cell by making a query to X.



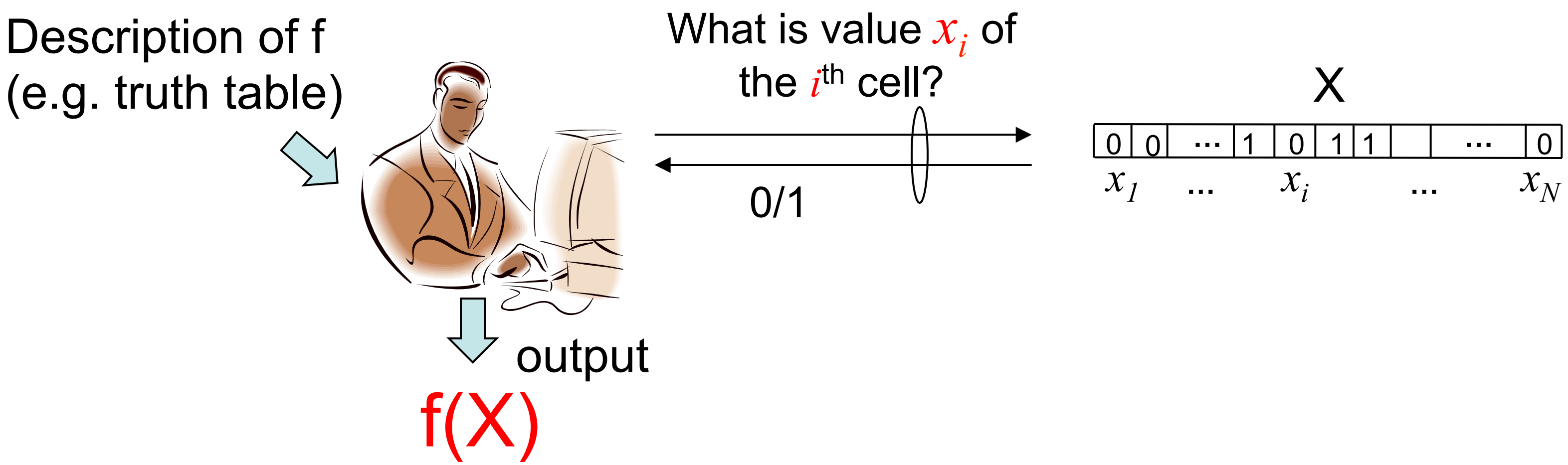

f(X)

- Cost measure:= # of queries to be made.
  (All other computation is free.)
- R(f): Query complexity of f

  := # of queries needed to compute f for the worst input X

# Oracle Computation Model

- Can know the value of one cell by making a query to X.




- Cost measure:= # of queries to be made.
  (All other computation is free.)

Bounded error

- R(f): Query complexity of f

with error probability < 1/3

:= # of queries needed to compute f for the worst input X

# Quantum Computation

Qubit: A unit of quantum information.

A quantum state $|\phi\rangle$ of one qubit:
a unit vector in 2-dimensional Hilbert space.

For an orthonormal basis $\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)=(|0\rangle,|1\rangle)$,
$|\phi\rangle=\alpha|0\rangle+\beta|1\rangle$ where $\alpha,\beta\in\mathbf{C}$ and $|\alpha|^2+|\beta|^2=1.$

A quantum state $|\varphi\rangle$ of n qubits:
a unit vector in $2^n$-dimensional Hilbert space.

$$|\varphi\rangle=\sum_{i=0}^{2^n-1}\alpha_i|i\rangle \quad \text{for orthonormal basis } \{|i\rangle\}_i.$$

Quantum operation: only unitary operation $H|\varphi\rangle\rightarrow|\varphi'\rangle$

# Oracle Computation Model (Quantum)

- A quantum query is a linear combination of classical queries.
- Can know a linear combination of the value of cells per query.

Description of f
(e.g. truth table)

$$\sum \alpha_i |i\rangle|0\rangle \rightarrow$$

$$\leftarrow \sum \alpha_i |i\rangle|x_i\rangle$$

$$\left(\sum_i |\alpha_i|^2 = 1\right)$$

X

| 0 | 0 | … | 1 | 0 | 1 | 1 | | … | 0 |
|---|---|---|---|---|---|---|---|---|---|

$x_1$ … $x_i$ … $x_N$

output

f(X)

- Q(f): (Bounded-error) Quantum query complexity of f

  := # of quantum queries needed to compute f with error probability < 1/3

  for the worst input X

# Fundamental Problems

- What is the quantum/classical query complexity of function f ?
- For what function f, is quantum computation faster than classical one?

In particular, Boolean functions are major targets.

This talk focuses on
Boolean functions in bounded-error setting
(constant error probability is allowed).

# Previous Works

- (Almost) No quantum speed up against classical.
  - PARITY, MAJORITY [BBCMdW01].
    - $\Omega(N)$ quantum queries are needed.

- Polynomial quantum speed up against classical
  - OR [Gro96], AND-OR trees [HMW03,ACRSZ07]
    - Quantum $O(\sqrt{N})$ v.s. Classical $\Omega(N)$.
  - k-threshold functions for k<< N/2 [BBCMdW01]
    - Quantum $\Theta(\sqrt{kN})$ v.s. Classical $\Omega(N)$.
  - Testing graph properties (N=n(n-1)/2 variables)
    - Triangle: Quantum $O(n^{1.3})$ [MSS05]
    - Star: Quantum $\Theta(n^{1.5})$ [BCdWZ99]
    - Connectivity: Quantum $\Theta(n^{1.5})$ [DHHM06]

    } Classical $\Omega(n^2)$

But much less is known except for the above typical cases.
→We investigate the query complexity of the families defined a natural parameter.

# On-set of Boolean Functions

We consider the *size of the on-set* of a Boolean function as a parameter.

*On-set* $S_f$ of a Boolean function f:

The set of input $X \in \{0,1\}^N$ for which $f(X)=1$.

Ex.)

On-set $S_f$ of $f=(x_1 \wedge x_2) \vee x_3$ :

$(x_1,x_2,x_3)=(1,1,0), (1,1,1), (0,0,1), (0,1,1), (1,0,1)$.

The size of $S_f$ is 5.

# Our Results (1/2)

$F_{N,M}$: family of $N$-variable Boolean functions $f$ whose on-set is of size $M$.

## Query complexity of the functions in $F_{N,M}$

$$(poly(N) \le M \le 2^{N^d} \text{ with } 0 < d < 1)$$

**Quantum Q(f)**

(1) Hardest case: $\Theta\left(\sqrt{N \dfrac{\log M}{\log N}}\right)$

(2) Easiest case: $\Theta\left(\sqrt{N}\right)$

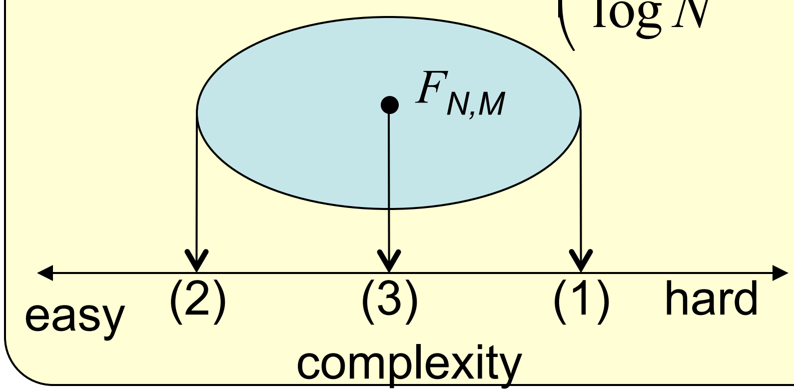
(3) Average Case: $O\left(\log M + \sqrt{N}\right)$,

$$\Omega\left(\frac{\log M}{\log N} + \sqrt{N}\right)$$

<<

**Classical R(f)**

$\Omega(N)$



# Our results (2/2)

## Our hardest-case complexity gives the tight complexity of some graph property testing.

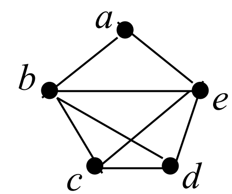
Does G have property P?

Is there edge (i,j) ?

Yes/No

$$\begin{array}{c|ccccc} & a & b & c & d & e \\ \hline a & 0 & 1 & 0 & 0 & 1 \\ b & 1 & 0 & 1 & 1 & 1 \\ c & 0 & 1 & 0 & 1 & 1 \\ d & 0 & 1 & 1 & 0 & 1 \\ e & 1 & 1 & 1 & 1 & 0 \end{array} =$$

Adjacency matrix of G




Unknown G

n(n-1)/2 variables

- (Planarity testing) Is G planar? : $Q(f)=\Theta(n^{1.5})$ . $R(f)=\Omega(n^2)$

  (For a given adjacency list, O(n) time complexity [Hopcroft-Tarjan74])

- (Graph Isomorphism testing) Is G isomorphic to a fixed graph G' ? :

  $Q(f)=\Theta(n^{1.5})$. ($R(f)=\Omega(n^2)$ [DHHM06])

By setting M = # of graphs with property P.

# OUTLINES OF PROOFS

# Our Results(1/2)

$F_{N,M}$: family of $N$-variable Boolean functions $f$ whose on-set is of size $M$.

## Query complexity of the functions in $F_{N,M}$

$(poly(N) \le M \le 2^{N^d}$ with $0 < d < 1)$

**Quantum Q(f)**

(1) Hardest case: $\Theta\left(\sqrt{N \frac{\log M}{\log N}}\right)$

(2) Easiest case: $\Theta\left(\sqrt{N}\right)$

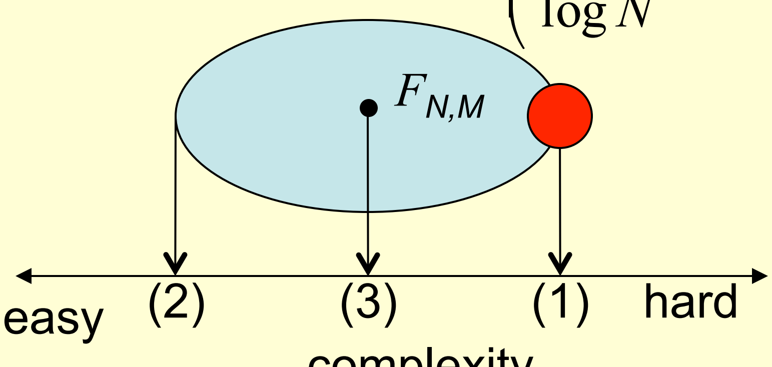
(3) Average Case: $O\left(\log M + \sqrt{N}\right)$,

$\Omega\left(\frac{\log M}{\log N} + \sqrt{N}\right)$

<< 

**Classical R(f)**

$\Omega(N)$



# Hardest-case Bound

Theorem : For any function $f \in F_{N,M}$,

$$Q(f) = \Theta\left(\sqrt{N \frac{\log M}{\log N}}\right)$$

if $poly(N) \le M \le 2^{N^d}$ for some constant $d\,(0 < d < 1)$.

Proof.

Lower Bound:

By showing a function for every $M$ which has $O\left(\sqrt{N \frac{\log M}{\log N}}\right)$ complexity.

(The function is similar to $t$ - threshold function for $t = \frac{\log M}{\log N}$.)

Upper bound:

Use the algorithm [AIKMRY07] for Oracle Identification Problem.

# Oracle Identification Problem (OIP)

- Given a set of M candidates, identify the N-bit string in the oracle. .

Oracle (N=8)

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $x_i$ | ? | ? | ? | ? | ? | ? | ? | ? |

Candidate Set (N=8, M=4)

Can see the contents w/o making queries.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Candidate 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Candidate 2 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| Candidate 3 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Candidate 4 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |

# Hardest-case Bound

## Proof (Continued)

Theorem[AIKMRY07]:

OIP can be solved with bounded-error by making $O\left(\sqrt{N\frac{\log M}{\log N}}\right)$ quantum queries,

if $poly(N) \le M \le 2^{N^d}$ for some constant $d\,(0 < d < 1)$.

Idea:

- Set the onset $S_f$ to the candidate set of OIP and run the algorithm for OIP to get an estimate $Y \in S_f$ of X.

- By definition, Y=X (with high probability) iff f(X)=1.

Test if X=Y,
which can be done with quantum query complexity O(√N). ■

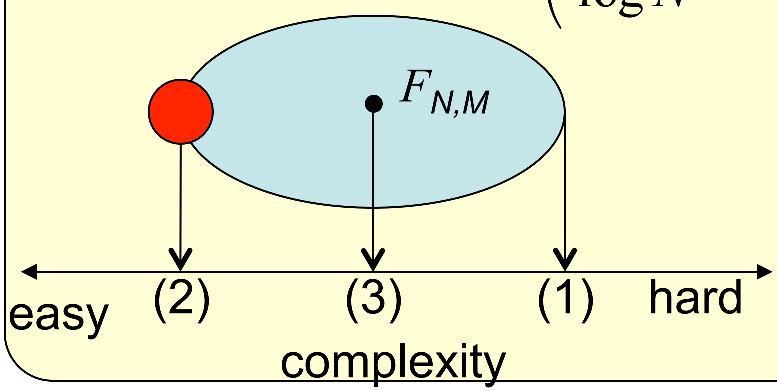
# Our Results (1/2)

$F_{N,M}$: family of $N$-variable Boolean functions $f$ whose on-set is of size $M$.

**Query complexity of the functions in $F_{N,M}$**

$(poly(N) \le M \le 2^{N^d}$ with $0 < d < 1)$

| Quantum Q(f) | | Classical R(f) |
|---|---|---|
| (1) Hardest case : $\Theta\left(\sqrt{N \frac{\log M}{\log N}}\right)$<br>(2) Easiest case : $\Theta\left(\sqrt{N}\right)$<br>(3) Average Case : $O\left(\log M + \sqrt{N}\right)$, $\Omega\left(\frac{\log M}{\log N} + \sqrt{N}\right)$ | << | $\Omega(N)$ |



# Easiest-case Bound

Theorem : If $M \le 2^{\frac{N}{2+\varepsilon}}$ for any positive constant $\varepsilon$,
$Q(f) = \Theta(\sqrt{N})$ for any $f \in F_{N,M}$.

Proof: Use sensitivity argument.

Th.[Beals et al. 2001] $Q(f) = \Omega(\sqrt{s(f)})$

Assuming s(f)=o(N), we can conclude
a contradiction by simply counting,

$$| f^{-1}(1) | > 2^{\frac{N}{2+\varepsilon}} \ge M$$

We can construct a function with such quantum query complexity. ■

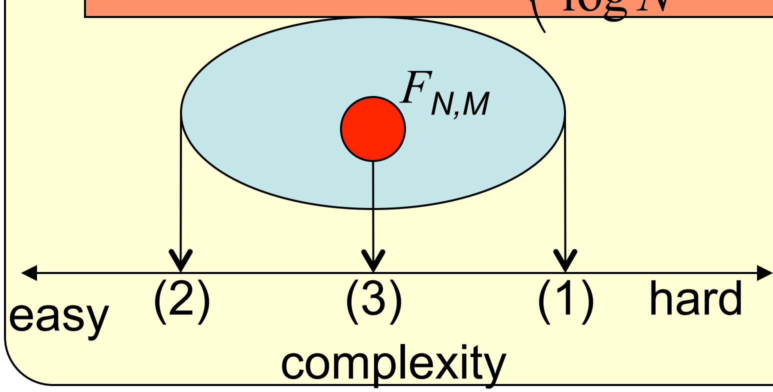
# Our Results (1/2)

$F_{N,M}$: family of $N$-variable Boolean functions $f$ whose on-set is of size $M$.

## Query complexity of the functions in $F_{N,M}$

$(poly(N) \leq M \leq 2^{N^d}$ with $0 < d < 1)$

| Quantum Q(f) | | Classical R(f) |
|---|---|---|
| (1) Hardest case : $\Theta\left(\sqrt{N \frac{\log M}{\log N}}\right)$ | | |
| (2) Easiest case : $\Theta\left(\sqrt{N}\right)$ | << | $\Omega(N)$ |
| (3) Average Case : $O\left(\log M + \sqrt{N}\right)$, $\Omega\left(\frac{\log M}{\log N} + \sqrt{N}\right)$ | | |

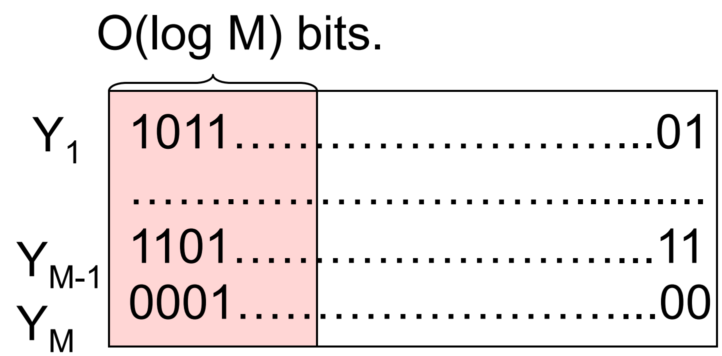


# Average-case Bound

Theorem: Average of $Q(f)$ over all $f \in F_{N,M}$ is $O(\log M + \sqrt{N})$.

Proof.

Claim: For almost all functions *f* in $F_{N,M}$, every element in the on-set $S_f$ differs from any other in the first O(log M) bits.




1. Make queries to the first O(log M) bits to identify a unique string Y in $S_f$
   (If there is no such Y, we are done: f(X)=0.)

2. Test if Y=X with O(√N) quantum queries.
   Y=X if and only if f(X)=1.

# Average-case Bound

Theorem : Average of $Q(f)$ over all $f \in F_{N,M}$ is $O\left(\frac{\log M}{c + \log N - \log\log M} + \sqrt{N}\right)$.

Proof

With one quantum query, $|\varphi_X\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (-1)^{x_i} |i\rangle$.

Claim: For almost all functions in $F_{N,M}$, every X, Y in the onset $S_f$ satisfy

$$\left|\langle \varphi_X | \varphi_Y \rangle\right| = \left| \frac{1}{N} \left(N - 2Ham(X,Y)\right) \right| > 2\sqrt{\frac{\log M}{N}}.$$

(Proof is by bounding Hamming distance with coding-theory argument and Chernoff-like bound.)

$\langle \varphi_X | \varphi_Y \rangle$ is large enough to identify $X$ in $S_f$ with

$O\left(\frac{\log M}{c + \log N - \log\log M}\right)$ copies of $|\varphi_X\rangle$

according to quantum state discrimination theorem [HW06]. ■

# Average-case Bound

$$\text{Theorem}: \text{Average of } Q(f) \text{ over all } f \in F_{N,M} \text{ is } \Omega(\log M / \log N + \sqrt{N}).$$

Actually, we prove stronger statement.

# Average-case Bound

Theorem : Average of unbounded - error query complexity over all $f \in F_{N,M}$ is $\Omega(\log M / \log N + \sqrt{N})$.

Unbounded-error: error probability is 1/2-ε for arbitrary small ε

Proof: Use the next theorem.

Theorem[Anthony1995 + Next Talk] The number of Boolean functions f whose unbounded query complexity is d/2 is

$$T(N,d) \le 2\sum_{k=0}^{D-1}\binom{2^N-1}{k} \text{ for } D = \sum_{i=0}^{d}\binom{N}{i}.$$

For $d = \dfrac{\log M}{2\log N}$, we can prove

$T\left(N, \dfrac{\log M}{2\log N}\right)$ is much smaller than $\dbinom{2^N}{M}$, i.e., the size of $F_{N,M}$.

# Our Quantum Complexity

$F_{N,M}$: family of $N$-variable Boolean functions $f$ whose on-set is of size $M$.

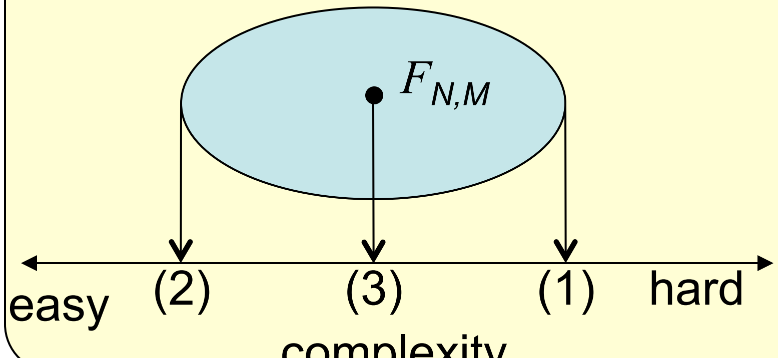
## Quantum query complexity of the functions in $F_{N,M}$

For $poly(N) \le M \le 2^{N^d}$ with $0 < d < 1$,

(1) Hardest case: $\Theta\left(\sqrt{N \frac{\log M}{\log N}}\right)$ ⇨ $\Theta\left(\sqrt{N \frac{\log M}{c + \log N - \log\log M}}\right)$ $\left(1 \le M \le 2^{N/(\log N)^{2+\varepsilon}}\right)$

(2) Easiest case: $\Theta\left(\sqrt{N}\right)$

(3) Average Case: $O\left(\log M + \sqrt{N}\right)$, $\Omega\left(\frac{\log M}{\log N} + \sqrt{N}\right)$ ⇨ $\Theta\left(\frac{\log M}{c + \log N - \log\log M} + \sqrt{N}\right)$ $\left(1 \le M \le 2^N / 2\right)$



# Application: Planarity Testing

Theorem:

$$\mathrm{R}(f_{\mathrm{planarity}}) = \Theta(n^{1.5}), \text{ while } \mathrm{R}(f_{\mathrm{planarity}}) = \Theta(n^2).$$

Proof.

Since the planar graph has at most 3n-6 edges.

$$M = (\#\text{ of planar graphs}) \leq \binom{\#\text{ of possible edges}}{3n-6} = \binom{n(n-1)/2}{3n-6} \leq 2^{6n\log n}$$

$$\text{By the hardest-case complexity, } \sqrt{N\frac{\log M}{\log N}}, \text{ we can obtain the upper bound.}$$

For the lower bound,

we carefully prepare a set of planar graphs and a set of non-planar graphs ,

and then apply the quantum/classical adversary method [Amb01,Aar04].

# Summary

- Proved the tight quantum query complexity of the family of Boolean functions with fixed on-set size M.
- Functions with on-set size M have various quantum query complexity, while their randomized query complexity is Ω(N) for $poly(N) \leq M \leq 2^{N^d}$.

  (For large M, the functions may have small randomized query complexity.)
- On-set size is a very simple and natural parameter, which enables us to easily analyze the query complexity of some Boolean functions with our bounds.
- In particular, we proved the tight quantum query complexity of some graph property testing problems.

Thank you!