# Optimizing Quantum Oracle for Standard Hash Algorithms

Korea University
Sangheon Lee, Mingyu Cho

December 13, 2020

# Table of Contents

# Introduction

# Previous Work

- Studied the basics of quantum computing and Grover's Algorithm
- Analysis of reversible quantum S-DES oracle and its construction
- Implementation of quantum S-DES oracle with Grover's Algorithm in Microsoft Q#

# Motivation

- Targeting SHA-2/3
  - Current de facto standard hash algorithm
- Goal is to perform a preimage attack.
  - The most straightforward to apply Grover's Algorithm.
  - It will be assumed that Grover's Algorithm(or some variation of it) will be applied
  - Therefore target is to optimize the quantum oracle for SHA-2/3.

# Recap: Grover's Algorithm

- Searching in an unordered database among $2^n$ elements takes $\tilde{\mathcal{O}}(2^{n/2})$ time complexity and $\tilde{\mathcal{O}}(n)$ quantum space complexity
- Proved to be optimal in general
- Specifically, improved Grover's Algorithm for collision search yields $\tilde{\mathcal{O}}(2^{n/3})$ time complexity and (quantum) space complexity [BHT98]
  - However, quantum queries are costly in this algorithm

# Chailloux's Algorithm [CNPS17]

- Use poly-time qubits and reduce time complexity to $\tilde{\mathcal{O}}(2^{2n/5})$ for collision search and $\tilde{\mathcal{O}}(2^{3n/7})$ for multi-target preimage attacks with additional classic memory

# SHA-2/3 Pre-image Attack Cost Estimation [ADMG+16]

- Suggest cost metric for quantum computation based on surface code
- Theoretically implement reversible SHA-2/3 quantum circuits
- Estimate required physical resources and its scale

# Expectations

- Focus on micro-scale improvement(regarding quantum oracle)
- May be able to reduce the cost in at least one of the following metrics:
  - The cost metric from before
  - The raw number of gates
  - The number of levels after achieving parallelization
- Main target will be the cost metric, not the gates or the levels

# In-depth Topics

- Surface code(Toric code) : topological quantum error correcting code
- Cost metric based on surface code
- T-par [AMM14] : an quantum circuit optimization tool
- Advanced quantum circuit (in-place adder, *etc.*)

# Surface Code

# Surface code and physical qubits

This summary is heavily based on [FMMC12].

- Surface code is a method to construct *logical* qubits from physical qubits with acceptable relativce error tolerance [CRSS97]
- Logical qubits are more efficient than their physical counterparts
- The tolerance of surface codes to errors is high as about 1%
- However, ensuring high tolerance requires massive physical qubits and sequential Toffoli gates
    - “We assume an error rate approximately one-tenth the threshold rate, which implies that we need about 14,500 physical qubits per logical qubit to give a sufficiently low logical error rate to successfully execute the algorithm”

# Surface code and physical qubits

- However, ensuring high tolerance requires massive physical qubits and sequential Toffoli gates
  - "A much larger part of the surface code is however needed to generate and purify the special ancilla $|A_L\rangle$ states that are used in the Toffoli gates."
  - Applying Shor's Algorithm to 2,000-bit integer requires $2.2 \times 10^{12}$ $|A_L\rangle$ states and takes about 26.7 hours
  - The surface code needs to generate these states in a timely manner

# Surface code and physical qubits

- However, ensuring high tolerance requires massive physical qubits and sequential Toffoli gates
  - "We assume an error rate approximately one-tenth the threshold rate, which implies that we need about 14,500 physical qubits per logical qubit to give a sufficiently low logical error rate to successfully execute the algorithm"
  - In result, 58 million qubits are required for computation
  - Additionally 1 billion qubits are required for generating $|A_L\rangle$ states
- Hopefully, "the size of the quantum computer is quite sensitive tothe error rate in the physical qubits."
  - "For example, improving the overall error rate by about a factor of ten, as detailed in Appendix M, can reduce the number of physical qubits by about an order of magnitude, to about 130 million, although leaving the execution time unchanged."

# Qubit operators

Quantum computation is based on qubits: two-level quantum systems

- Based on quantum physics and election spins

These electron spins can be reprensented by various operators such as Pauli-X, Y, Z operators.

# Qubit operators

Basic recap of Pauli operators and qubit operators:

- Ground state for $\hat{Z}$ axis : $|g\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- Excited state : $|e\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $\hat{Z} = \hat{\sigma_z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, with eigenvalues $+1$ and $-1$ for $|g\rangle$, $|e\rangle$.
- $\hat{X} = \hat{\sigma_x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, with eigenvalues $+1$ and $-1$ for
  $|+\rangle = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle)$,
  $|-\rangle = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) = \frac{1}{\sqrt{2}}(|g\rangle - |e\rangle)$.
- $\hat{Y} = -i\hat{\sigma}_y = \hat{Z}\hat{X} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, which is real unlike $\hat{\sigma}_y$.

# Qubit operators

These qubit operators satisfy the following:

$$\begin{aligned} \hat{X}^2 &= -\hat{Y}^2 = \hat{Z}^2 = I \\ \hat{X}\hat{Z} &= -\hat{Z}\hat{X} \\ [\hat{X}, \hat{Y}] &= \hat{X}\hat{Y} - \hat{Y}\hat{X} = -2\hat{Z} \end{aligned} \tag{1}$$

Also note that each measurement based on each quantum operators yields one of its eigenstates. For example, $M_Z$ will return $|g\rangle$ or $|e\rangle$, while $M_X$ will return $|+\rangle$ or $|-\rangle$.

# Qubit operators

- Any two-level quantum system that satisfies (1) can be used as a qubit.
- There are infinitely many quantum states...
- Hopefully, Solovay–Kitaev theorem states that an arbitrary quantum gate can be approximated by a reasonably short sequence of universal gates.
    - An $m$ constant qubit gate can be approximated to an $\epsilon$-error by $\mathcal{O}(m \log^c(m/\epsilon))$ for some $c$.
- One of the common quantum universal gate set is 'Clifford + T' gate set: CNOT, $\hat{H}$(Hadamard gate), $\hat{S}$ and $\hat{T}$.[1]

[1] $T^4 = S^2 = Z = R_\pi$, where $R_\theta = \left(\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right)$.

# Single-qubit errors

Qubits errors are random $\hat{X}$ bit-flip and/or $\hat{Z}$ phase-flip.

- These single-qubit errors can be undone after detection
  - Erroneous $\hat{Z}$ can be cancelled with another $\hat{Z}$ since $\hat{Z}^2 = I$
- Also, applying qubit operation does not alter the probability of its eigenstate

Thus detection, rather than correction, is the key point of surface code.

# Single-qubit errors

- However since $[\hat{X}, \hat{Z}] \neq \hat{0}$, sequential measurements of $\hat{X}$ and $\hat{Z}$ might conflict each other (since $\hat{X}\hat{Z} \neq \hat{Z}\hat{X}$).
- This problem can be avoided by measuring multi-qubits at once. Consider qubit $a$ and $b$ and operation $\hat{X}_a\hat{X}_b$ and $\hat{Z}_a\hat{Z}_b$, then these two operations do commute!

$$
\begin{aligned}
[\hat{X}_a\hat{X}_b, \hat{Z}_a\hat{Z}_b] &= (\hat{X}_a\hat{X}_b)(\hat{Z}_a\hat{Z}_b) - (\hat{Z}_a\hat{Z}_b)(\hat{X}_a\hat{X}_b) \\
&= \hat{X}_a\hat{Z}_a\hat{X}_b\hat{Z}_b - \hat{Z}_a\hat{X}_a\hat{Z}_b\hat{X}_b \\
&= (-\hat{Z}_a\hat{X}_a)(-\hat{Z}_b\hat{X}_b) - (\hat{Z}_a\hat{X}_a)(\hat{Z}_b\hat{X}_b) \\
&= \hat{0}
\end{aligned} \tag{2}
$$

Therefore $\hat{X}_a\hat{X}_b$ and $\hat{Z}_a\hat{Z}_b$ can be used as basis of 2-qubit system.

# Single-qubit errors

In this case, the following are the (entangled) eigenstates, which does not change its state after measurements by $\hat{X}_a\hat{X}_b$ or $\hat{Z}_a\hat{Z}_b$.

| $\hat{Z}_a\hat{Z}_b$ | $\hat{X}_a\hat{X}_b$ | $\lvert\psi\rangle$ |
|---|---|---|
| $+1$ | $+1$ | $(\lvert gg\rangle + \lvert ee\rangle)/\sqrt{2}$ |
| $+1$ | $-1$ | $(\lvert gg\rangle - \lvert ee\rangle)/\sqrt{2}$ |
| $-1$ | $+1$ | $(\lvert ge\rangle + \lvert eg\rangle)/\sqrt{2}$ |
| $-1$ | $-1$ | $(\lvert ge\rangle - \lvert eg\rangle)/\sqrt{2}$ |

For example, if $\lvert\psi\rangle = (\lvert gg\rangle - \lvert ee\rangle)/\sqrt{2}$, then $\hat{Z}_a\hat{Z}_b\lvert\psi\rangle = \lvert\psi\rangle$ and $\hat{X}_a\hat{X}_b\lvert\psi\rangle = -\lvert\psi\rangle$ since $(+1, -1)$ are its eigenvalues.

# Single-qubit errors

If there is an random $\hat{X}$ or $\hat{Z}$ error on a single qubit, the original eigenstate will change into another eigenstate.

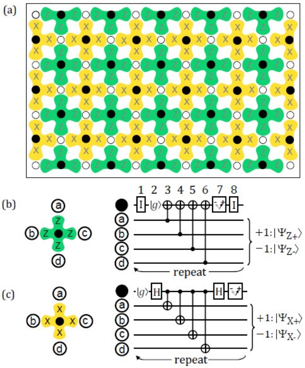
- If $|\psi\rangle = (|gg\rangle - |ee\rangle)/\sqrt{2}$, then $\hat{X}_a\,|psi\rangle = -(|ge\rangle - |eg\rangle)/\sqrt{2}$, where eigenvalues changed to $(-1, -1)$.
- However, $\hat{X}_b\,|psi\rangle$ also yields $-(|ge\rangle - |eg\rangle)/\sqrt{2}$
- Errors can be detected, but cannot be uniqlely identified
- A simple form of error detection code

Operator products such as $\hat{X}_a\hat{X}_b$ or $\hat{Z}_a\hat{Z}_b$ are called stabilizers

- Eigenstates are measureable without any quantum information loss

# Surface code structure

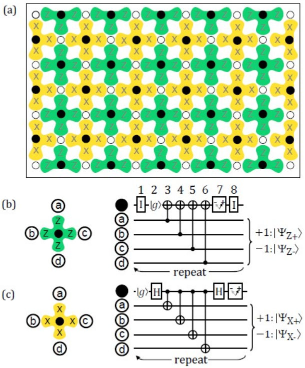
The following figure is the implementation of the surface code on a two-dimensional array of physical qubits.




- The open circles are *data qubits* where the computational quantum states are stored.
- The filled circles are *measurement qubits*.
- These qubits should be able to perform initialization, single-qubit rotation, CNOT between neighbors, SWAP and $M_Z$.

# Surface code structure

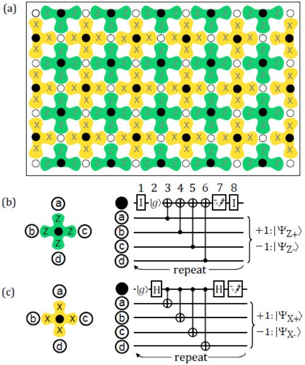
The following figure is the implementation of the surface code on a two-dimensional array of physical qubits.




- Measurement qubits are used to stabilize and alter the quantum state of data qubits
    - Two types exist: "measure-Z" qubits (colored in green) and "measure-X" qubits (colored in yellow)
- Each data qubits are connected to two measure-Z qubits and measure-X qubits

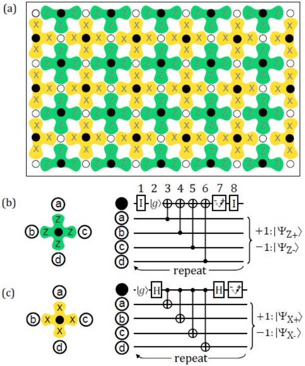
# Surface code structure

The following figure is the implementation of the surface code on a two-dimensional array of physical qubits.




- Measurement qubits either measures $\hat{X}$ or $\hat{Z}$ stabilizer (i.e., $\hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d$ or $\hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d$)
- Figure (b) and (c) represents the quantum circuit of stabilizer measurement
- For any adjacent data qubits with state $|\psi\rangle$, $\hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d\,|\psi\rangle = X_{abcd}\,|\psi\rangle$ holds.
- After measurement, the cycle is repeated (to negate the effects)

# Surface code structure

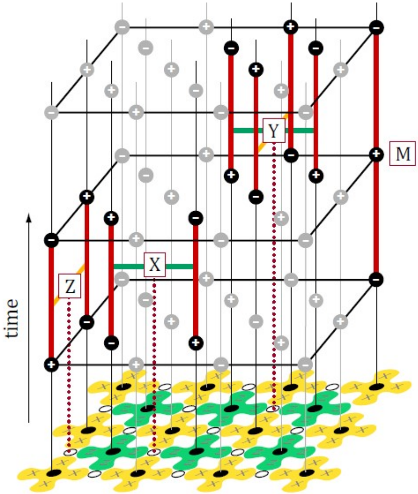



- The techniques used to apply $\hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d$ and $\hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d$ are somewhat standard, but is surely crucial.
- On (b), identity matrix is included in measure-Z to minimize error and match timing with measure-X.

# Quinscent state

- The forementions state $|\psi\rangle$ is not the ground state $|g\rangle$
- Instead it is determined by its eigenvectors
- The data qubits are projected to these states
- They are not fully entangled to all of the qubits, but are locally entangled

# Single-qubit error in a nutshell



# Single phase-flip error

Consider a random $\hat{Z}$ phase-flip error $\hat{I}_a + \epsilon\hat{Z}_a$ where $|\epsilon| \ll 1$. Let $|\psi'\rangle \to (\hat{I}_a + \epsilon\hat{Z}_a)\,|\psi\rangle$, then $|\psi'\rangle = \hat{Z}_a\,|\psi\rangle$ with probability $|\epsilon|^2$.

$$\begin{aligned} \hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d(\hat{Z}_a\,|\psi\rangle) &= -\hat{Z}_a(\hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d\,|\psi\rangle) \\ &= -\hat{Z}_a X_{abcd}\,|\psi\rangle \end{aligned} \tag{3}$$

Thus, a single $\hat{Z}$ phase-flip error can be detected by measure-X with any info loss since $|\psi\rangle$ is the eigenstate of stabilizers $\hat{X}_a\hat{X}_b\hat{X}_c\hat{X}_d$ and $\hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d$. However, it will not change the result of measure-Z.

$$\begin{aligned} \hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d(\hat{Z}_a\,|\psi\rangle) &= \hat{Z}_a(\hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d\,|\psi\rangle) \\ &= \hat{Z}_a Z_{abcd}\,|\psi\rangle \end{aligned} \tag{4}$$

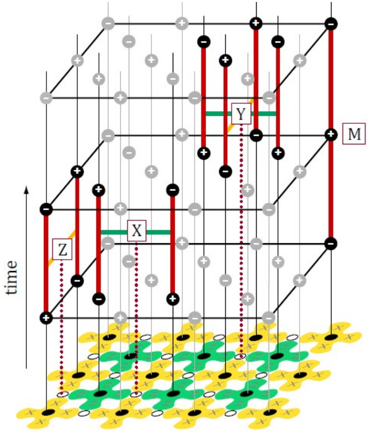
# Other errors

- We can apply $\hat{Z}_a$ to negate this error, but applying another quantum operator may yield other errors
- Marking that the qubit is phase-flipped by software is better
- Similarly, a single $\hat{X}$ bit-flip error can be detected by measure-Z.

Measurement can also yield errors, which has to be considered.

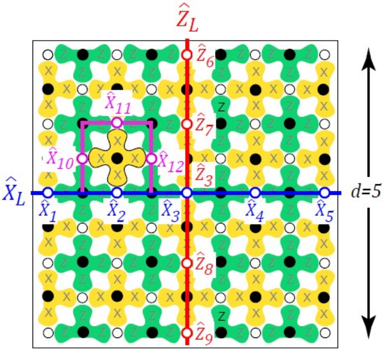
- However, it is very unlikely to happen twice in a row or more
- Even so, several measurements are required to ensure the result

# Single-qubit error in a nutshell (recap)



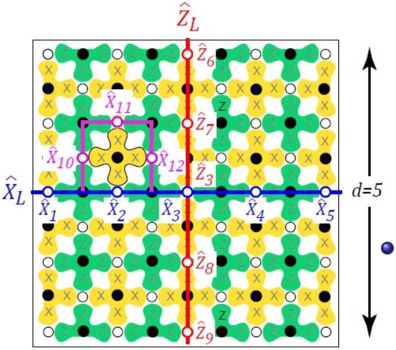

Vertical heavy lines indicate each errors

# Logical operators in a nutshell




$\hat{X}_L$ and $\hat{Z}_L$ can manipulate qubits without altering the measurement results.

# 9 by 9 surface code

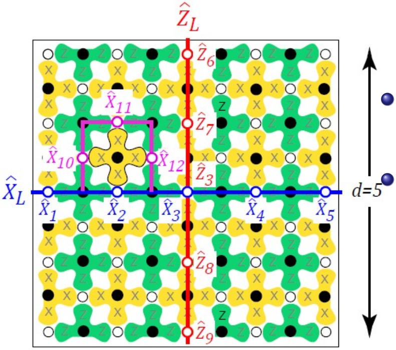



- This 9 by 9 surface code contains 41 data qubits and 40 measurement qubits
  - Thus there are $2 \times 41$ degrees of freedom and $2 \times 40$ constraints with the stabilizer
  - All of these constraints are linearly independent
- Where are the remaining two unconstrained degrees of freedom? How can we manipulate those?

# 9 by 9 surface code

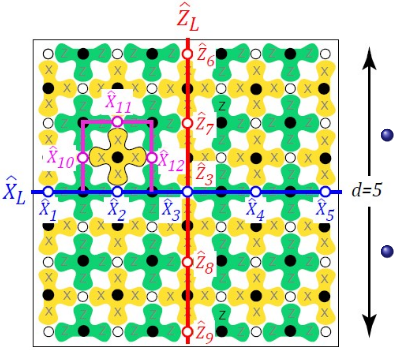
- A single data qubit $\hat{X}$ operation only alters the measure-Z outcomes on either side
  - As mentioned, all $\hat{X}_i$ operations mutually commute.
- However, two simultaneous $\hat{X}$ operations on two data qubits $a$ and $b$ that both neighbor the same measure-$Z$ qubit commutes with the $\hat{Z}$ stabilizer
- As a formula, $[\hat{Z}_a\hat{Z}_b\hat{Z}_c\hat{Z}_d, \hat{X}_a\hat{X}_b] = 0$.

# 9 by 9 surface code




- We first apply $\hat{X}_1$. To negate the measure-Z, we also apply $\hat{X}_2$, and so on.
- Thus, $\hat{X}_L = \hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4\hat{X}_5$ commutes with any stabilizer.
- $|\psi_X\rangle = \hat{X}_L |\psi\rangle$ is different from $|\psi\rangle$, but its measurement will yield $|\psi\rangle$.

# 9 by 9 surface code




- Since $\hat{X}_L$ cannot be written as a product of stabilizers, this operator manipulates one of the unconstrained degrees of freedom.
- Similarly $\hat{Z}_L = \hat{X}_6\hat{X}_7\hat{X}_3\hat{X}_8\hat{X}_9$ manipulates the other degree of freedom.
- Are there any other methods to manipulate unconstrained degrees of freedom other than $\hat{X}_L$ and $\hat{Z}_L$? No!

# Clifford + T gate

# Why Clifford + T?

- By Gottesman-Knill Theorem, Clifford gates can be simulated efficiently on a classical computer.
- T-gate is eqal to $R_{\pi/4} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{i} \end{pmatrix}$, a $\frac{\pi}{4}$ rotation.
- Clifford+T gate utilizes only one inefficient gate, while is universal.

# Ripple-Carry Adder

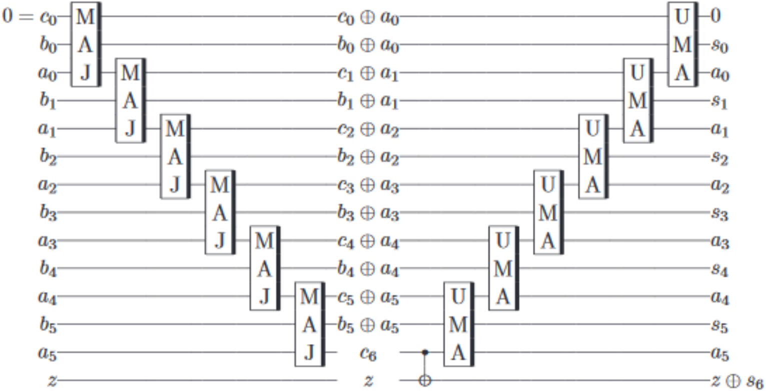



Figure 4: A simple ripple-carry adder for $n = 6$.

# References

# References I


- Matthew Amy, Olivia Di Matteo, Vlad Gheorghiu, Michele Mosca, Alex Parent, and John Schanck, *Estimating the cost of generic quantum pre-image attacks on sha-2 and sha-3*, International Conference on Selected Areas in Cryptography, Springer, 2016, pp. 317–337.
- Matthew Amy, Dmitri Maslov, and Michele Mosca, *Polynomial-time t-depth optimization of clifford+ t circuits via matroid partitioning*, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **33** (2014), no. 10, 1476–1489.
- Gilles Brassard, Peter Høyer, and Alain Tapp, *Quantum cryptanalysis of hash and claw-free functions*, Latin American Symposium on Theoretical Informatics, Springer, 1998, pp. 163–169.

# References II

André Chailloux, María Naya-Plasencia, and André Schrottenloher, *An efficient quantum collision search algorithm and implications on symmetric cryptography*, International Conference on the Theory and Application of Cryptology and Information Security, Springer, 2017, pp. 211–240.

A Robert Calderbank, Eric M Rains, Peter W Shor, and Neil JA Sloane, *Quantum error correction and orthogonal geometry*, Physical Review Letters **78** (1997), no. 3, 405.

Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, Physical Review A **86** (2012), no. 3.