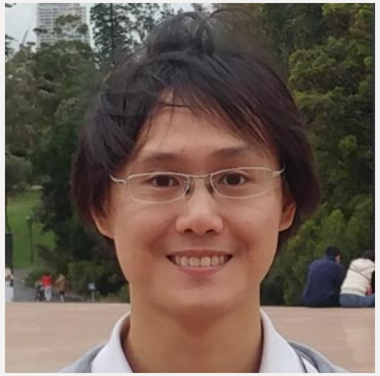
A unified architecture for quantum lookup tables

2406.18030

Shuchen Zhu



Aarthi Sundaram
Microsoft Quantum

Guang Hao Low
Google Quantum

Duke

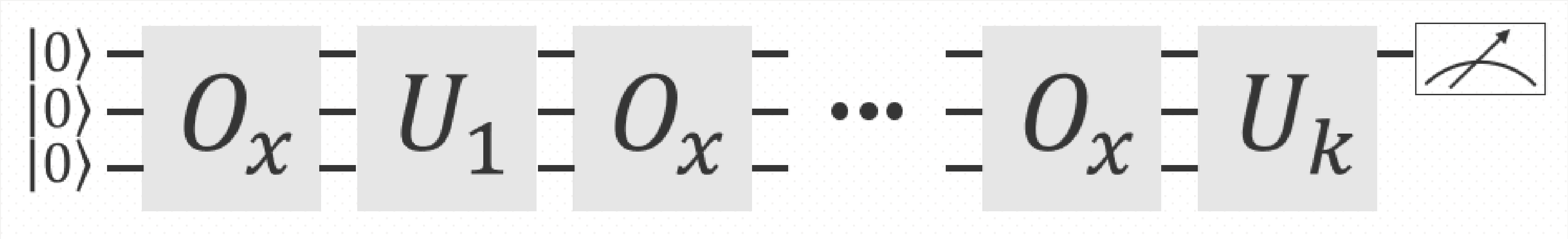
What is quantum lookup table

- It stores classical data and allows queries to be made in superposition
- It is a general-purpose architecture for the implementation of unstructured quantum oracles

$$|i\rangle|b\rangle \longrightarrow \boxed{O_x} \longrightarrow |i\rangle|b \cdot x_i\rangle$$

Why quantum lookup table?

- Most quantum algorithms are naturally phrased as oracle query algorithms

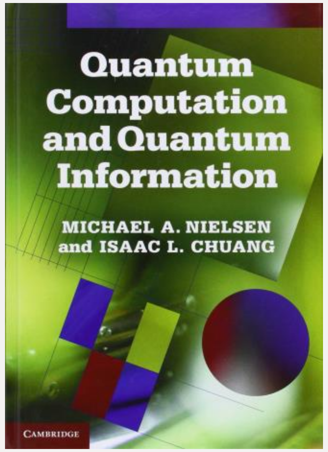


- Oracle provides description of Hamiltonian in quantum simulation
- Oracle encodes classical datasets into quantum state in quantum machine learning

- Essential to achieve quantum computational advantage

- Quantum state preparation is expensive; the overhead may negate the quantum advantage

Prior arts



Fan-out

PHYSICAL REVIEW A **78**, 052310 (2008)

Architectures for a quantum random access memory

Vittorio Giovannetti,[1] Seth Lloyd,[2] and Lorenzo Maccone[3]
[1]*NEST CNR-INFM & Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126, Pisa, Italy*
[2]*MIT, Research Laboratory of Electronics and Department of Mechanical Engineering, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*
[3]*QUIT—Quantum Information Theory Group, Dipartimento di Fisica "A. Volta," Università di Pavia, via A. Bassi 6, I-27100 Pavia, Italy*
(Received 7 August 2008; published 5 November 2008)

QRAM

⟨Quantum
the open journal for quantum science
PAPERS PERSPECTIVE

Trading T gates for dirty qubits in state preparation and unitary synthesis

Guang Hao Low[1,2], Vadym Kliuchnikov[1,2], and Luke Schaeffer[1,3,4]
[1]Quantum Architectures and Computation, Microsoft Research, Washington, Redmond, USA
[2]Azure Quantum, Microsoft, Washington, Redmond, USA
[3]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[4]Joint Center for Quantum Information and Computer Science, University of Maryland, Maryland, College Park, USA

SELECT-SWAP

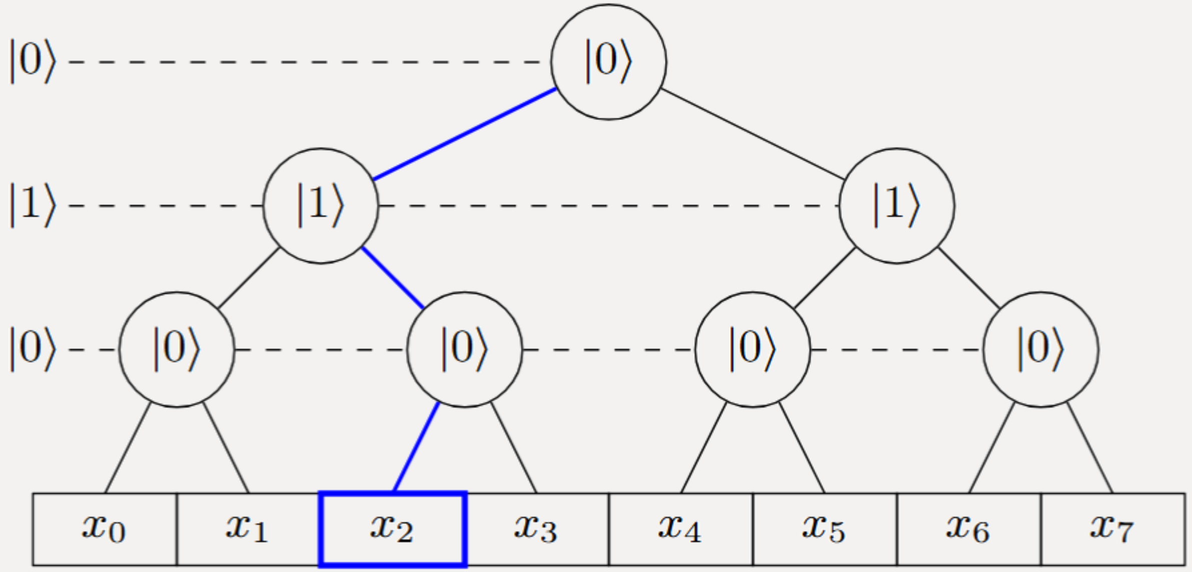
PHYSICAL REVIEW X **8**, 041015 (2018)

Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity

Ryan Babbush,[1,*] Craig Gidney,[2] Dominic W. Berry,[3] Nathan Wiebe,[4] Jarrod McClean,[1] Alexandru Paler,[5] Austin Fowler,[2] and Hartmut Neven[1]
[1]*Google Inc., Venice, California 90291, USA*
[2]*Google Inc., Santa Barbara, California 93117, USA*
[3]*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*
[4]*Microsoft Research, Redmond, Washington 98052, USA*
[5]*Institute for Integrated Circuits, Linz Institute of Technology, 4040 Linz, Austria*

QROM

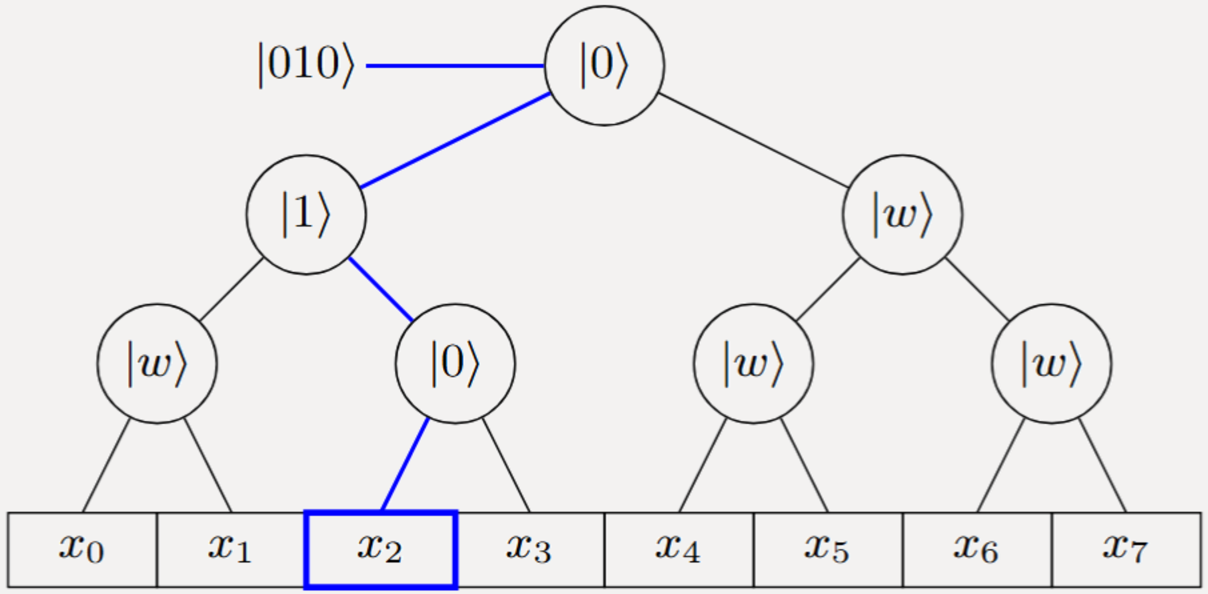
Fan-out architecture



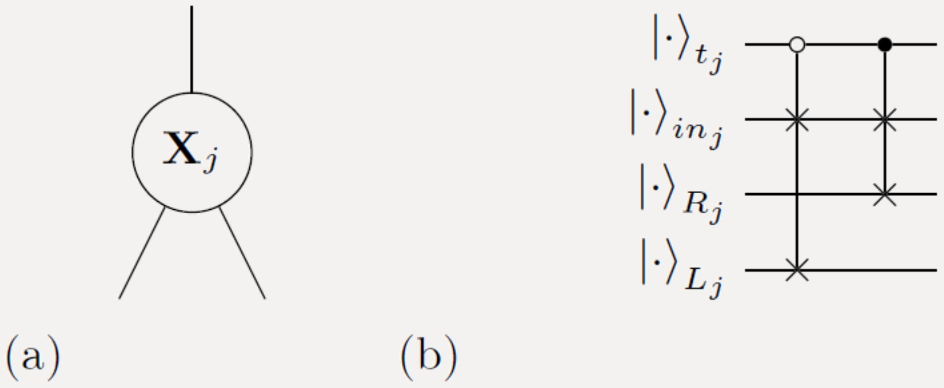
- Linear infidelity

Nielsen, Michael A., and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

Bucket-brigade (QRAM) architecture



CSWAP router

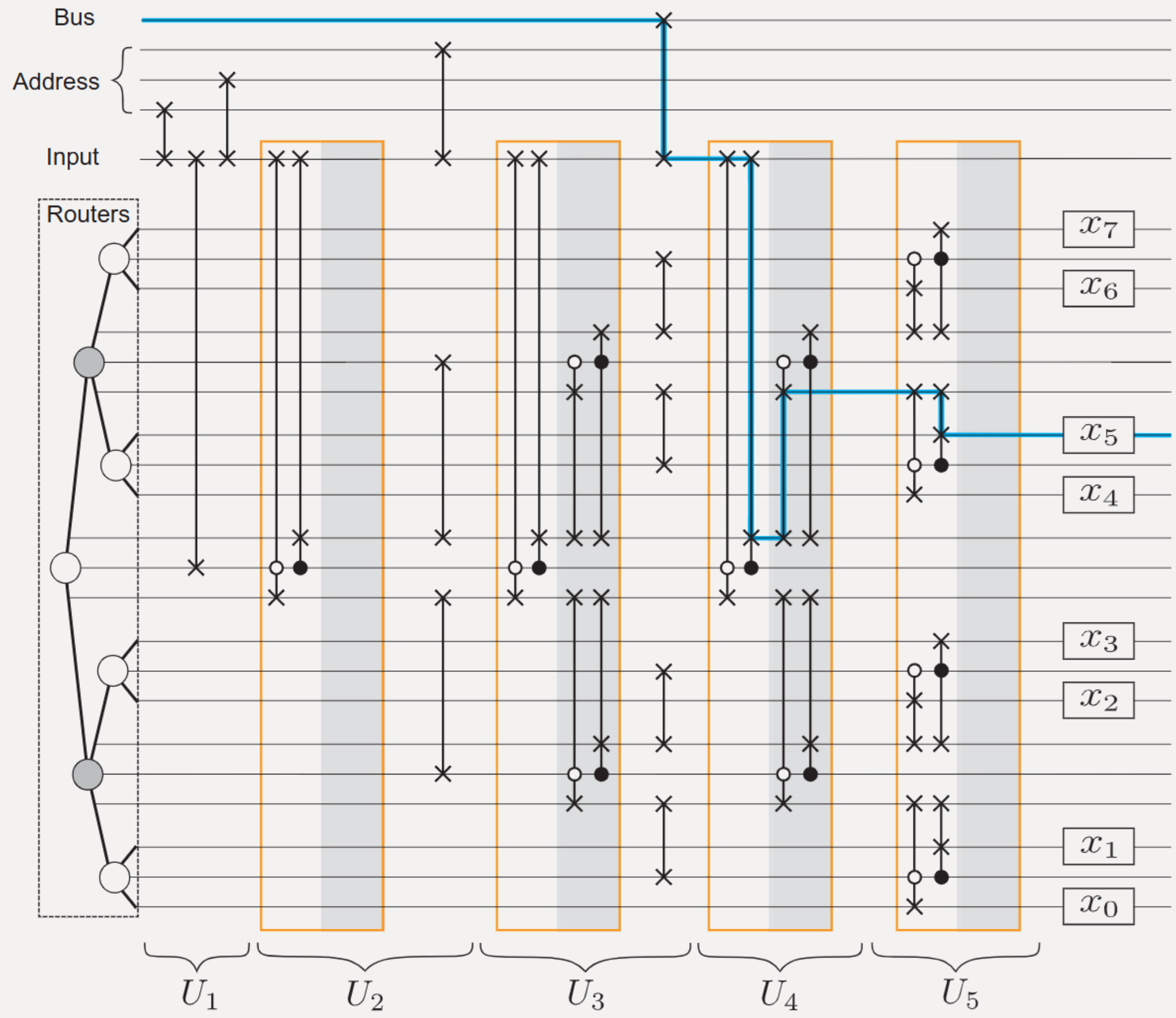


- Low query time $O(\log N)$
- Low infidelity $O(\log^2 N)$
- High T-gate count: $O(N)$

Giovannetti, Vittorio, Seth Lloyd, and Lorenzo Maccone. "Architectures for a quantum random access memory." *Physical Review A* 78, no. 5 (2008): 052310.
Hann, Connor T., Gideon Lee, S. M. Girvin, and Liang Jiang. "Resilience of quantum random access memory to generic noise." PRX Quantum 2, no. 2 (2021): 020311.

Duke

Bucket-brigade (QRAM) circuit

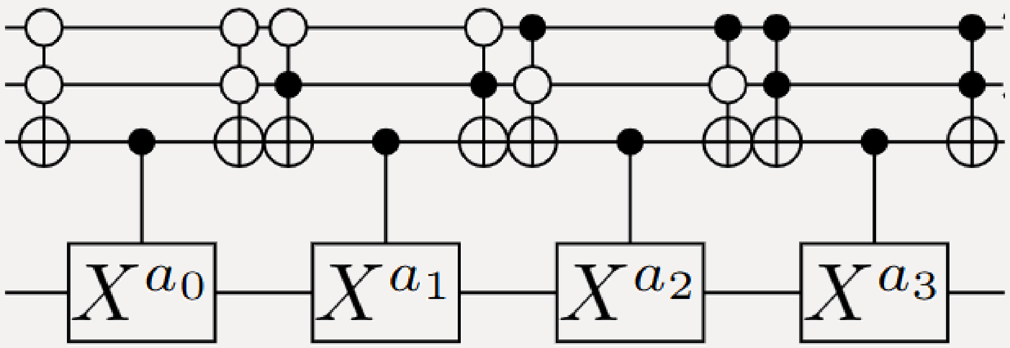


- All-to-all connectivity assumption
- Assume same gate error

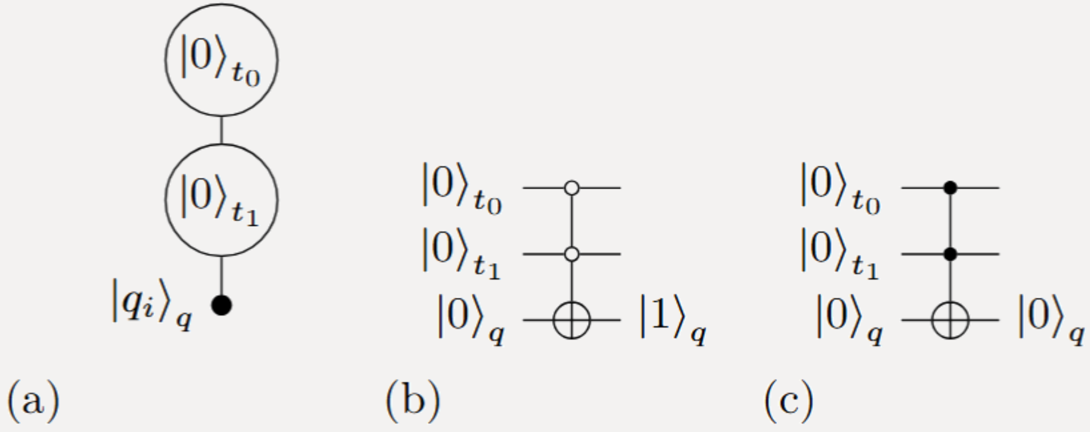
Duke

Hann, Connor T., Gideon Lee, S. M. Girvin, and Liang Jiang. "Resilience of quantum random access memory to generic noise." PRX Quantum 2, no. 2 (2021): 020311.

QROM architecture



Linear router

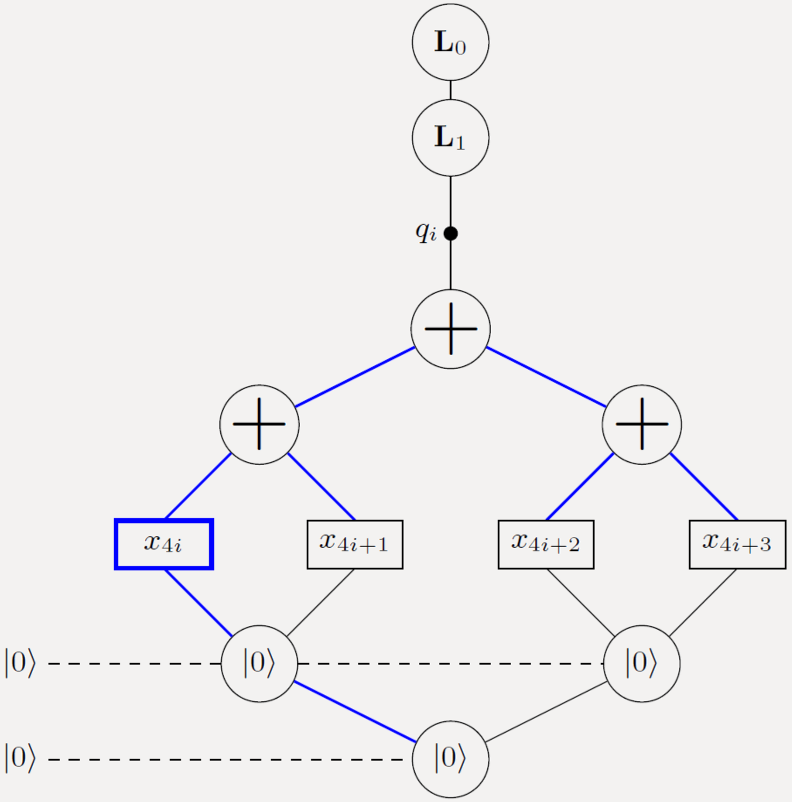


- Low qubit count $O(\log N)$
- High query time $O(N)$
- High infidelity $O(N)$
- High T-gate count: $O(N)$

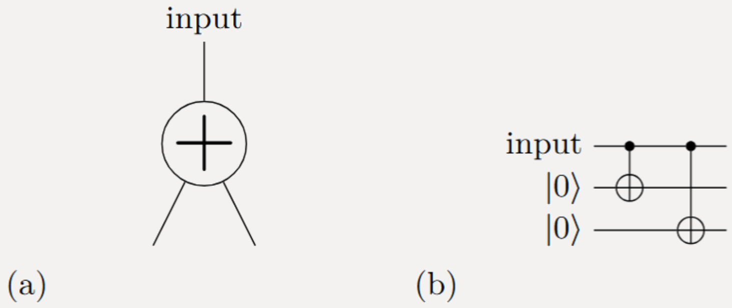
Babbush, Ryan, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven. "Encoding electronic spectra in quantum circuits with linear T complexity." *Physical Review X* 8, no. 4 (2018): 041015.

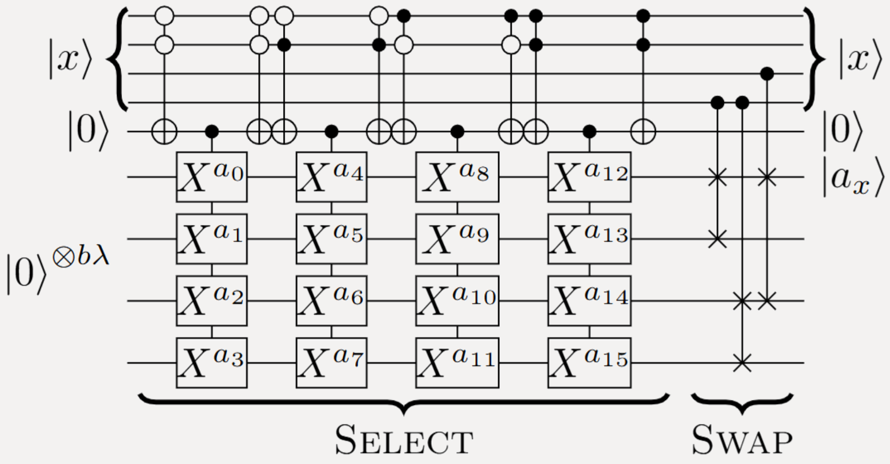
Duke

SELECT-SWAP architecture



CNOT router





- Sublinear qubit count $O(\sqrt{N})$
- Sublinear query time $O(\sqrt{N})$
- Fixed high infidelity $O(N)$
- Sublinear T-gate count: $O(\sqrt{N})$

Low, Guang Hao, Vadym Kliuchnikov, and Luke Schaeffer. "Trading T gates for dirty qubits in state preparation and unitary synthesis." *Quantum* 8 (2024): 1375.

Duke

Summary of prior art

Architecture	Infidelity	Query-depth	Qubits	T-gates	Layout
QRAM [19, 24]	$O(\log^2 N)$	$O(\log N)$	$O(N)$	$O(N)$	all-to-all
QROM [4]	$O(bN)$	$O(N)$	$O(\log N + b)$	$O(N)$	all-to-all [a]
SELECT-SWAP [23]	$O(bN)$	$O(\frac{N}{\lambda} + \log \lambda)$	$O(\log N + b\lambda)$	$O(\frac{N}{\lambda} + b\lambda)$	all-to-all

- Can we drop the all-to-all assumption in the analysis?
- Can we preserve the good properties from all these frameworks?

Duke

Our result

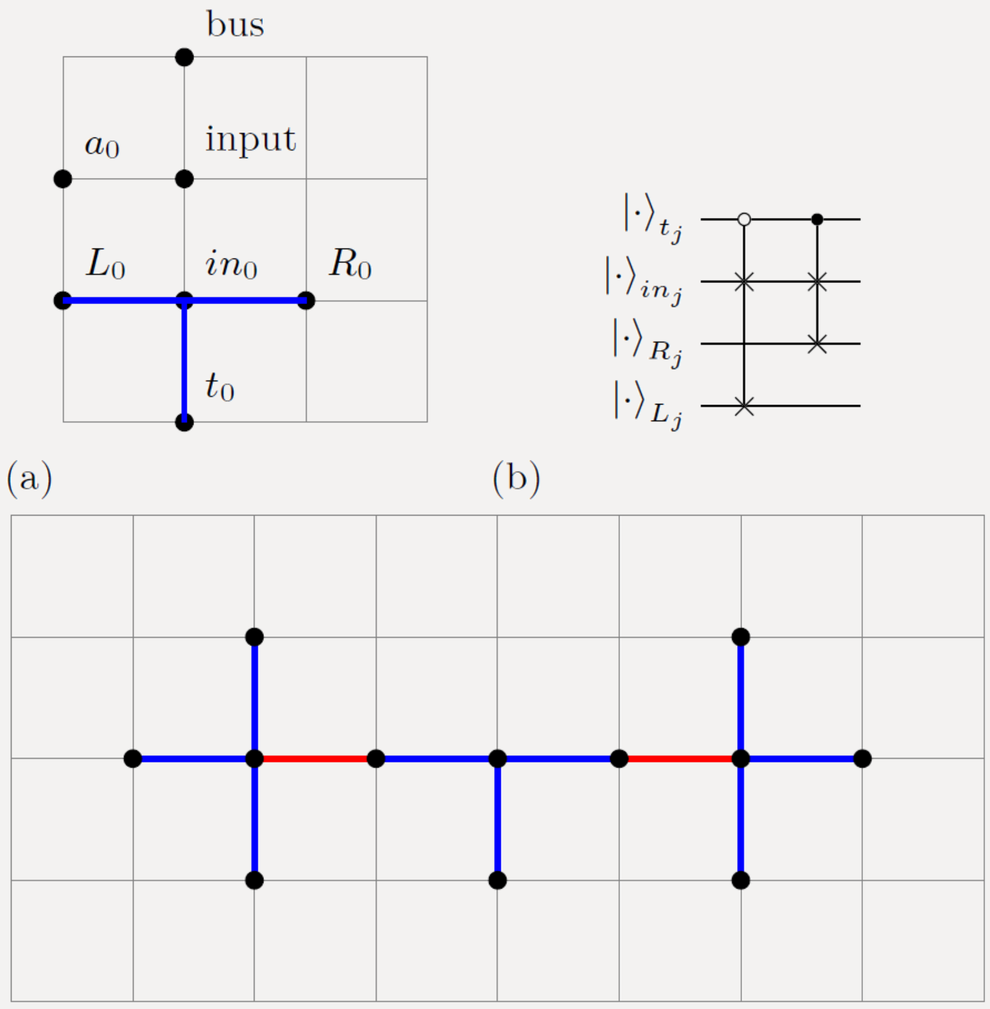
Architecture	Infidelity	Query-depth	Qubits	T-gates	Layout
QRAM [19, 24]	$O(\log^2 N)$	$O(\log N)$	$O(N)$	$O(N)$	all-to-all
QROM [4]	$O(bN)$	$O(N)$	$O(\log N + b)$	$O(N)$	all-to-all [a]
SELECT-SWAP [23]	$O(bN)$	$O(\frac{N}{\lambda} + \log \lambda)$	$O(\log N + b\lambda)$	$O(\frac{N}{\lambda} + b\lambda)$	all-to-all
general single-bit (Sec. III)	$\tilde{O}(\frac{\gamma N}{\lambda})$	$O(\frac{N}{\lambda} \log N)$	$O(\log \frac{N}{\lambda} + \lambda)$	$O(\frac{N}{\gamma} + \frac{N}{\lambda} \log \frac{N}{\lambda} + \lambda)$	local, planar
general multi-bit parallel (Sec. V)	$\tilde{O}(\frac{b\gamma N}{\lambda})$	$O(\frac{N}{\lambda}(\log bN))$	$O(\log \frac{N}{\lambda} + b\lambda)$	$O(\frac{N}{\lambda} \log \frac{N}{\lambda} + \frac{bN}{\gamma} + b\lambda)$	local, planar
general multi-bit sequential (Sec. V)	$\tilde{O}(\frac{b\gamma N}{\lambda} + b^2)$	$O(\frac{N}{\lambda} \log(bN) + b)$	$O(\log \frac{N}{\lambda} + b\lambda)$	$O(\frac{N}{\gamma} + \frac{N}{\lambda} \log \frac{N}{\lambda} + b\lambda)$	local, planar
single-bit (Sec. III)	$\tilde{O}(N^{3/4})$	$O(\sqrt{N} \log N)$	$O(\sqrt{N})$	$O(N^{3/4})$	local, planar
parallel multi-bit (Sec. V)	$\tilde{O}(bN^{3/4})$	$O(\sqrt{N} \log N)$	$O(b\sqrt{N})$	$O(bN^{3/4})$	local, planar
sequential multi-bit (Sec. V)	$\tilde{O}(bN^{3/4} + b^2)$	$O(\sqrt{N} \log(bN) + b)$	$O(b\sqrt{N})$	$O(N^{3/4} + b\sqrt{N})$	local, planar

- Planar layout
- Sublinear scalings for local connectivity
- Fine-tuned error dependence
- Unified framework
- Extend to large word size

Error type	Error rate symbol
Idling	ε_I
Qubit	ε_Q
Long range	ε_L
SWAP gate	ε_s
CSWAP gate	ε_{cs}
CNOT gate	ε_c
CCNOT gate	ε_{cc}

Duke

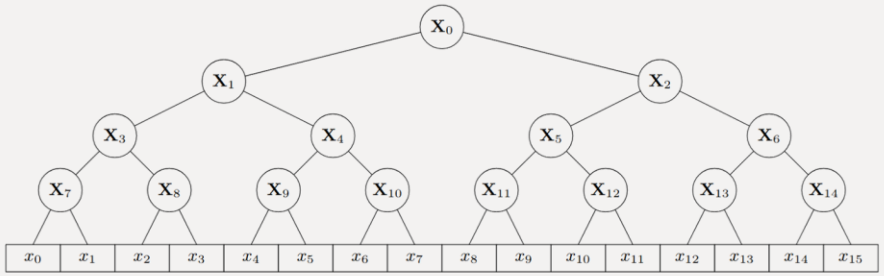
Planar layout for single CSWAP router

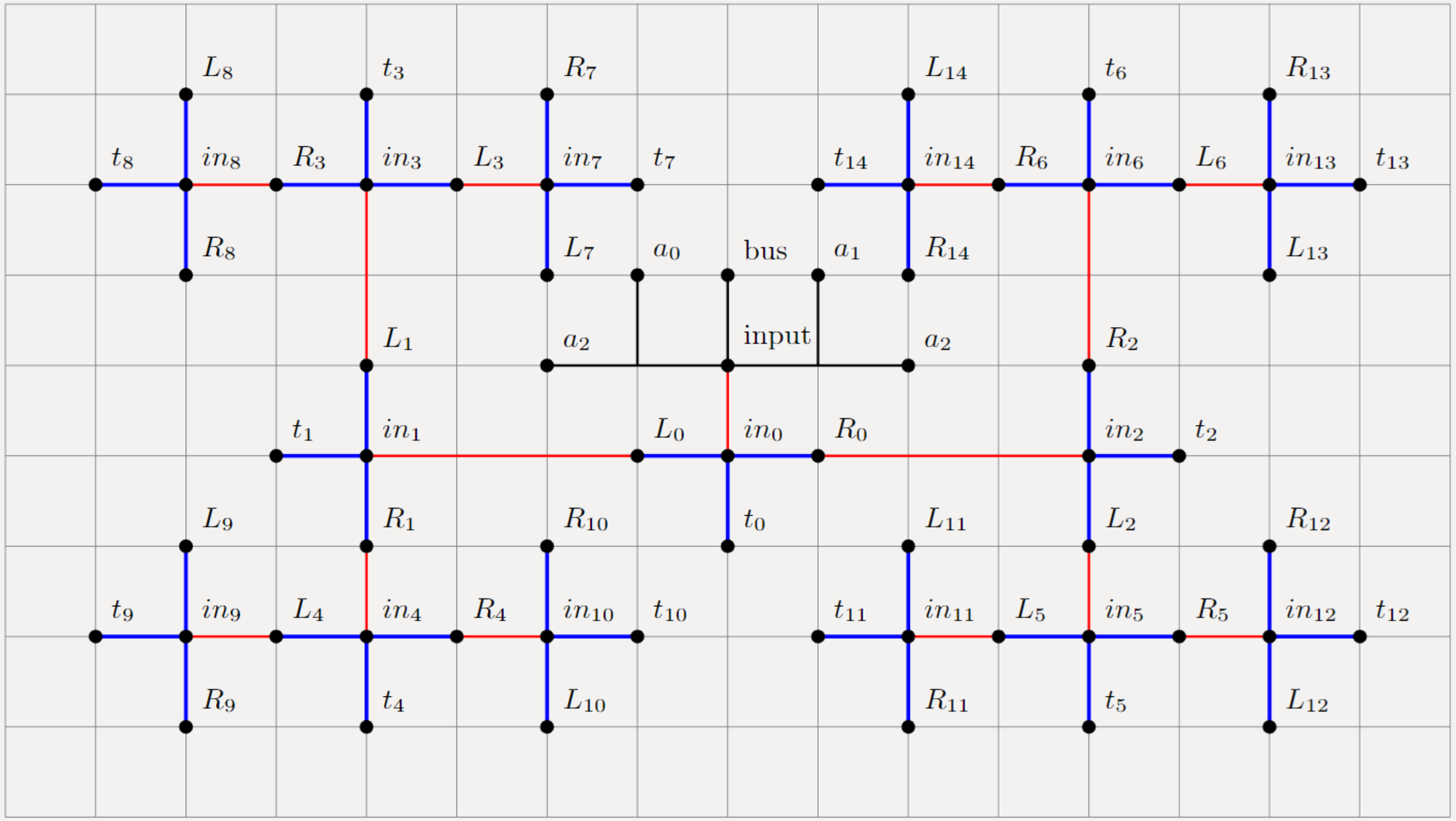


(c)

Duke

Planar layout for QRAM

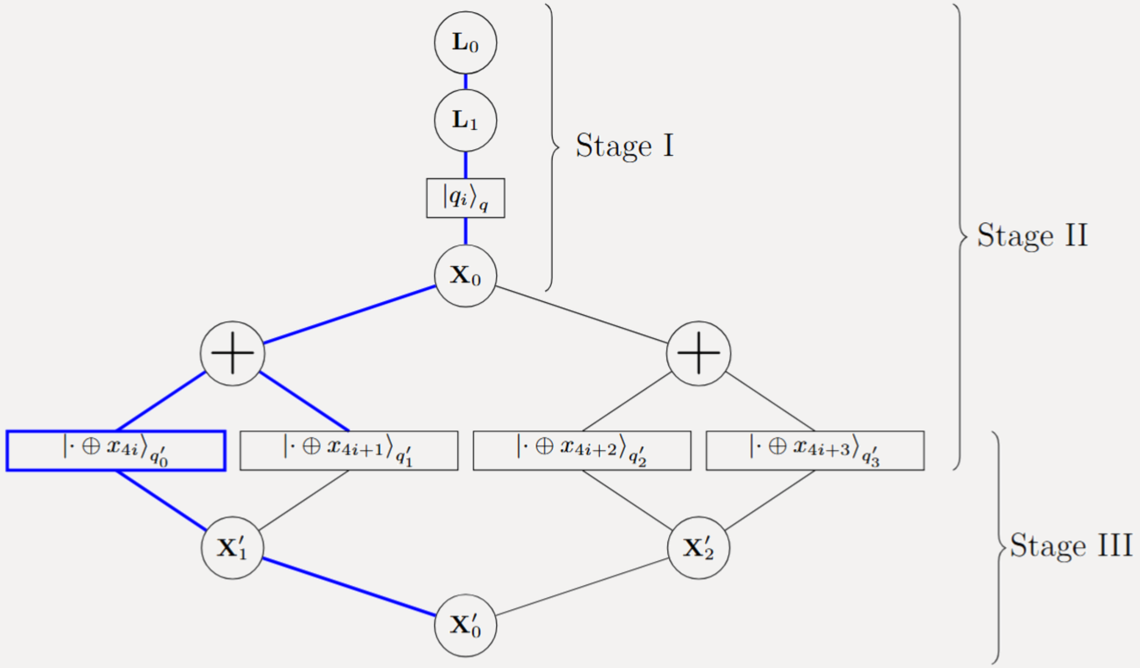




- This layout can be extended to our unified framework.

Duke

Ideal construction



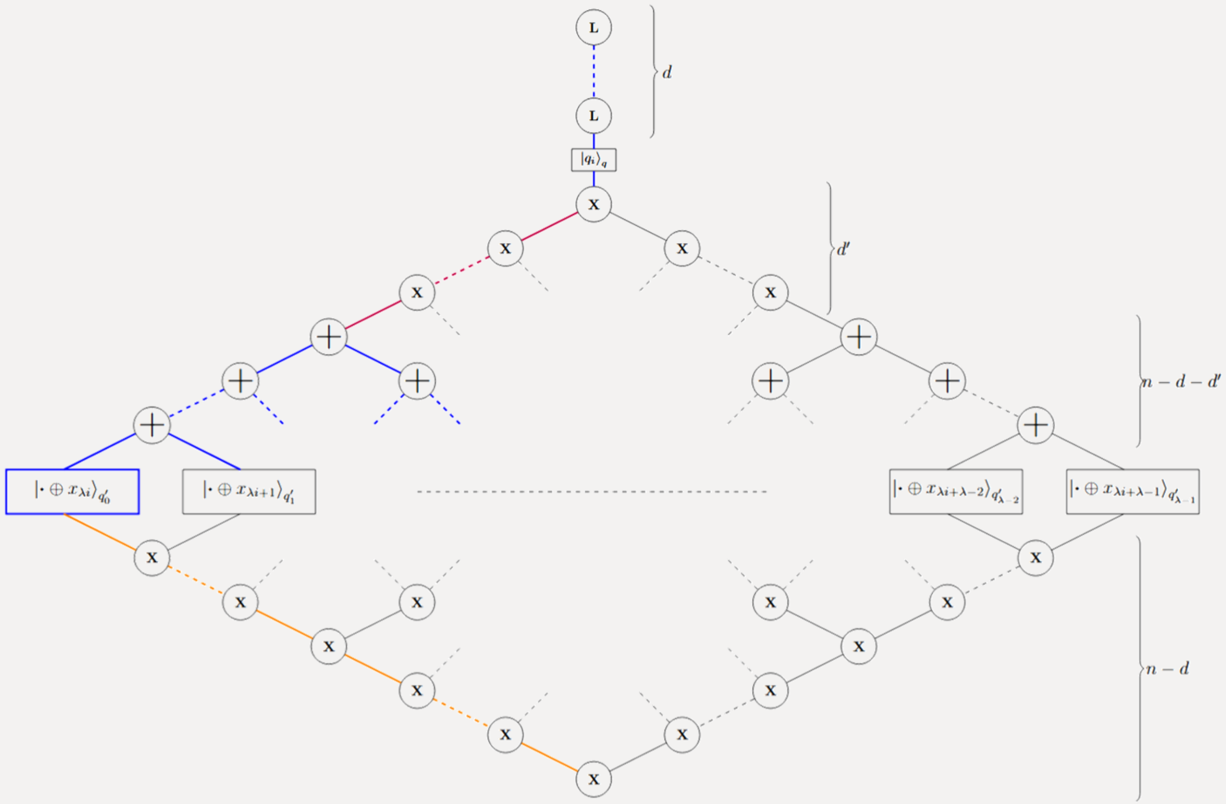
Algorithm 1 Pseudocode for quantum data lookup on memory of size N, partition size λ, and CNOT tree size γ. This corresponds to the high-level routing scheme shown in Figure 8. SetRouter(a, r) sets the address bit(s) a to router(s) of type $r \in \{\mathbf{X}, \mathbf{L}\}$ as described in Section II. For each SetRouter operation, a path connected by its corresponding routers is formed. RouteData(d, p) moves data qubit d from one end to the other end of the path p. We use $\tilde{p}$ to denote the reverse of the path p.

Setup: classical database `data` of size N with $n = \log_2 N$, memory partition $\lambda \leq N$ with $d = \log_2\left(\frac{N}{\lambda}\right)$ and CNOT tree size $\gamma \leq \lambda$ with $d' = \log_2\left(\frac{\lambda}{\gamma}\right)$.
Input: state $|a\rangle_{\text{add}}$ where $a = a_0 \dots a_{n-1}$ is an address of length n.
Output: state $|a\rangle_{\text{add}} |\texttt{data}[a]\rangle_{\text{out}}$ where $\texttt{data}[a]$ refers to the data at address a.

Stage I:
Initialize all registers of the quantum lookup table to $|0\rangle$.
SetRouter$(a_0 \dots a_{d-1}, \mathbf{L})$ ▷ Set d linear routers $\mathbf{L}_k$
SetRouter$(a_d \dots a_{d+d'-1}, \mathbf{X})$ ▷ Set d' CSWAP routers $\mathbf{X}$ in the corresponding path p_1 of the CSWAP tree
Stage II:
for i in $0 \dots 2^d - 1$ **do**
  if $i = a_0 \dots a_{d-1}$ **then** ▷ The state of Linear routers $\mathbf{L}$ determines the activation qubit q_i at round i
    Set $|q_i\rangle_q \leftarrow |1\rangle$
  else
    Set $|q_i\rangle_q \leftarrow |0\rangle$
  RouteData(q_i, p_1) ▷ Route q_i along the path p_1 determined by the CSWAP routers to a CNOT router
  Propagate q_i's signal to the γ leaves $|c_0, \dots, c_{\gamma-1}\rangle_{q'}$ of the CNOT tree.
  for j in $0 \dots \lambda - 1$ **do** ▷ Set the q' registers determined by q_i and `data`
    Apply unitary $U_{\lambda \cdot i+j}$ where $U_{\lambda \cdot i+j} = \text{CNOT}|c_j\rangle\,|q'_j\rangle$ if $\texttt{data}[\lambda \cdot i + j] = 1$ and $\mathbf{I}$ otherwise.
  Reset all qubits c_j and those in the CNOT tree to 0.
  RouteData$(q_i, \tilde{p}_1)$ ▷ Route q_i back up the CSWAP routers to its original location
  Reset all the qubits of the CSWAP routers except the router status qubits and q_i to 0.
Stage III:
SetRouter$(a_d \dots a_{n-1}, \mathbf{X})$ ▷ Set the status of $\log_2(\lambda)$ CSWAP in path p_2 routers based on the remaining address bits
RouteData(q'_ℓ, p_2), $\ell = a_d \dots a_{n-1}$ ▷ Route q'_ℓ along the query path p_2 set by the CSWAP routers to the output register

Duke

Main result



Theorem III.1. *Consider the quantum data lookup structure with the high-level scheme in Fig. 8 with N memory locations. Let $n = \log N$, $\lambda = 2^{n-d}$ be the partition size and $\gamma = 2^{n-d-d'}$ be the size of a CNOT tree with $d' \leq d \leq n$. The infidelity of this circuit is*

$$O\Bigg(\varepsilon_L\left(\frac{\gamma N}{\lambda}+\frac{N}{\lambda}\log\frac{\lambda}{\gamma}\right)+\varepsilon_s\log\frac{\lambda^2}{\gamma} + \varepsilon_I\left(\frac{N}{\lambda}\log N\left(\log\frac{N}{\gamma}+\gamma+\log\frac{\lambda}{\gamma}\right)+\text{polylog}\,\lambda\right) + \varepsilon_c\frac{\gamma N}{\lambda}+\varepsilon_{cc}\frac{N}{\lambda}\log\frac{N}{\lambda} + \varepsilon_{cs}\left(\frac{N}{\lambda}\log\frac{\lambda}{\gamma}+\log^2\frac{\lambda}{\gamma}+\log^2\lambda\right)\Bigg). \tag{3}$$

Moreover, the T count for this design is $O(\frac{N}{\gamma}+\frac{N}{\lambda}\log\frac{N}{\lambda}+\lambda)$, and its qubit count is $O(\log\frac{N}{\lambda}+\lambda)$.

Current progress in the field

Demonstrating Coherent Quantum Routers for Bucket-Brigade Quantum Random Access Memory on a Superconducting Processor

Sheng Zhang,[1,2,3] Yun-Jie Wang,[4] Peng Wang,[1,2,3] Ren-Ze Zhao,[1,2] Xiao-Yan Yang,[1,2] Ze-An Zhao,[1,2] Tian-Le Wang,[1,2] Hai-Feng Zhang,[1,2] Zhi-Fei Li,[1,2] Yuan Wu,[1,2] Hao-Ran Tao,[5] Liang-Liang Guo,[5] Lei Du,[5] Chi Zhang,[5] Zhi-Long Jia,[5] Wei-Cheng Kong,[5] Zhuo-Zhi Zhang,[1,2,3] Xiang-Xiang Song,[1,2,3] Yu-Chun Wu,[1,2,6] Zhao-Yun Chen,[6,*] Peng Duan,[1,2,†] and Guo-Ping Guo[1,2,5,‡]

[1] *Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui, 230026, China*
[2] *CAS Center For Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui, 230026, China*
[3] *Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, Jiangsu 215123, China*
[4] *Institute of the Advanced Technology, University of Science and Technology of China, Hefei, Anhui, 230088, China*
[5] *Origin Quantum, Hefei, Anhui 230088, China*
[6] *Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui, 230088, China*
(Dated: June 1, 2025)

A distillation–teleportation protocol for fault-tolerant QRAM

Alexander M. Dalzell,[1] András Gilyén,[2] Connor T. Hann,[1] Sam McArdle,[1] Grant Salton,[1,3] Quynh T. Nguyen,[4] Aleksander Kubica,[1,5] Fernando G.S.L. Brandão[1,6]

[1] *AWS Center for Quantum Computing, Pasadena, CA, USA*
[2] *HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary*
[3] *Amazon Quantum Solutions Lab, Seattle, WA, USA*
[4] *School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA*
[5] *Yale Quantum Institute & Department of Applied Physics, New Haven, CT, USA*
[6] *Institute for Quantum Information and Matter, Caltech, Pasadena, CA, USA*

Productionizing Quantum Mass Production

William J. Huggins,[1,*] Tanuj Khattar,[1] and Nathan Wiebe[2,3,4]

[1] *Google Quantum AI, Mountain View, CA, USA*
[2] *Department of Physics, University of Toronto, Toronto, ON, Canada*
[3] *Department of Computer Science, University of Toronto, Toronto, ON, Canada*
[4] *Pacific Northwest National Laboratory, Richland, WA, USA*
(Dated: June 6, 2025)

- How to benefit from qutrit/qudit.
- How to do QRAM simulation on a real machine to gain better intuition.
- If the distillation-teleportation's bottleneck is classical, can we improve it to make it work?

- Dream: $O(\log N)$ infidelity for planar layout.