

A resource-efficient quantum-walker Quantum RAM

Giuseppe De Riso¹, Giuseppe Catalano², Seth Lloyd³, Vittorio Giovannetti², Dario De Santis¹

¹Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126, Pisa, Italy

²NEST-CNR Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126, Pisa, Italy

³Dept. of Mech. Eng., Massachusetts Institute of Technology, 77 Mass. Av., Cambridge, MA 02139, USA

Paper available at <https://arxiv.org/pdf/2508.02855>

Introduction and motivation

A quantum Random Access Memory (qRAM) enables coherent access to a classical database in superposition. Consider a database M with $N = 2^n$ memory cells, where each cell $M^{(a)}$ is identified by a unique n -bit binary address $\mathbf{a} := a_1, a_2, \dots, a_n$ and stores m bits of classical information $\mathbf{b}^{(a)} := b_1^{(a)}, b_2^{(a)}, \dots, b_m^{(a)}$. When the address register A is prepared in a superposition $\sum_{\mathbf{a}} \alpha_{\mathbf{a}} |\mathbf{a}\rangle_A$ and $|0\rangle_D$ is a fiducial state of a data register D , the qRAM implements the unitary transformation:

$$\sum_{\mathbf{a}} \alpha_{\mathbf{a}} |\mathbf{a}\rangle_A |0\rangle_D \xrightarrow{\text{qRAM}} \sum_{\mathbf{a}} \alpha_{\mathbf{a}} |\mathbf{a}\rangle_A |\mathbf{b}^{(a)}\rangle_D . \quad (1)$$

Efficient qRAM is the bottleneck for quantum speedups in linear algebra [1–4], quantum chemistry [5–8], machine learning [9, 10], and cryptography [11, 12]. Yet, turning qRAM from a theoretical oracle into a physical device has proven daunting: proposed architectures often balance optimal query complexity with impractical physical requirements. While the “bucket-brigade” architecture [13] offered a seminal leap in noise resilience [14], its implementation remains challenging due to the active routing requirements. Recently, a quantum walker-based architecture [15] emerged as a promising alternative, theoretically achieving optimal circuit depth. However, this walker proposal suffers from a critical physical drawback: to ensure time efficiency, it requires long-range connectivity scaling with the inter-tree distance, along with a massive spatial overhead. We present a novel walker-based qRAM architecture [16] that achieves polylogarithmic access time in the database size N using exclusively *local, nearest-neighbor* interactions. By introducing a modular “backup walker” protocol, we decompose the routing logic into local collision-based gates. This approach allows us to condense the required hardware onto a *single binary tree* structure, matching the spatial footprint of the standard bucket brigade, thereby avoiding the massive overhead of parallel trees or long-range connections required by previous walker models. Crucially, our architecture operates with *passive*¹ dynamics: routing is determined solely by the internal quantum state of the walkers via unitary evolution, without requiring external control signals at each node. This feature satisfies a fundamental scalability criterion recently highlighted by Jaques and Rattew [17], who argue that the prohibitive overhead of “active” routing risks negating the asymptotic speedups of qRAM-based algorithms. By combining the noise resilience of the bucket brigade with the passive nature of quantum walkers, our model offers a realistic blueprint for the first practical large-scale qRAM.

Prior works and the gap we fill

The seminal “bucket-brigade” (BB) qRAM [13] introduced a routing scheme on a binary tree of depth n , where $N = 2^n$ leaves represent memory cells. Each node hosts a qutrit router that, upon interaction with address carriers, dynamically opens a path to the target. While this architecture ensures that only

¹Meaning that no address-dependent reconfiguration of the nodes is required.

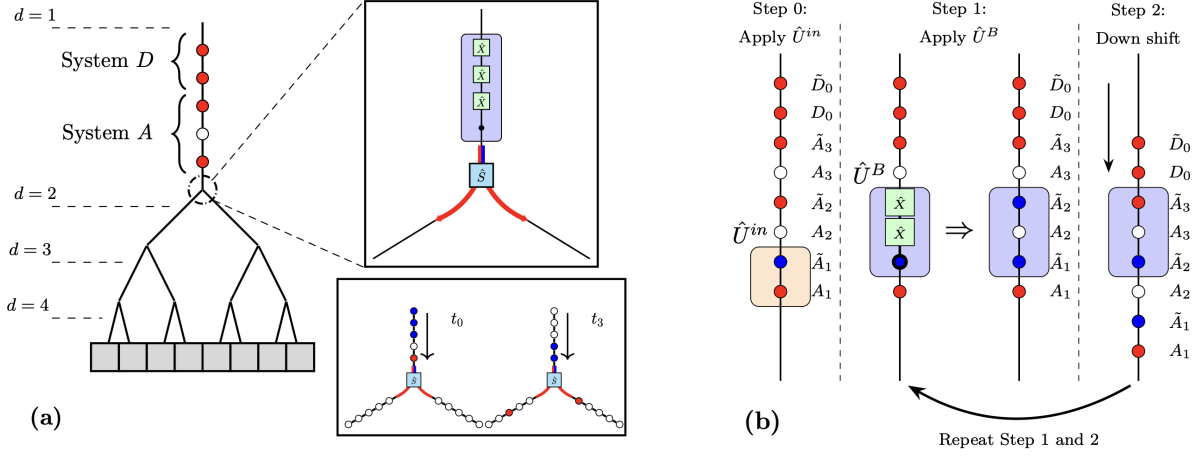


Figure 1: **(a)** Schematic of the proposed qRAM architecture on a single binary tree of depth $d = 3$. At the root node, $n = 3$ address walkers and $m + 1$ data walkers are injected sequentially. The upper-right inset shows the two operations at each node: the gate $\hat{U}^{(i)}$ (purple box), which conditionally flips the color of all subsequent walkers, and the scattering gate $\hat{S}$, which routes red walkers left and blue walkers right. The lower-right inset illustrates the scattering dynamics at two different time steps. **(b)** The backup-walker decomposition of $\hat{U}^{(1)}$ for address $a_1 = 1, a_2 = 0, a_3 = 0$. Step 0: $\hat{U}^{in}$ transfers the address bit from A_1 to its backup $\tilde{A}_1$. Steps 1–2 (iterated): the three-body gate $\hat{U}_B$ propagates the routing information to the next pair of walkers, followed by a spatial shift. The entire decomposition uses only nearest-neighbor Clifford gates.

$\mathcal{O}(\log N)$ nodes are active per query, granting high noise resilience [14], it requires $\mathcal{O}(N)$ three-level systems to be present in the hardware, even though only $\mathcal{O}(\log N)$ of them are activated per query. The need to maintain coherence and control over these active routing elements constitutes a substantial overhead. More recently, Asaka, Sakai, and Yahagi (ASY) proposed a walker-based model [15] that eliminates active routers entirely: routing is encoded in the walkers’ internal states, making the tree nodes passive scatterers. While theoretically promising, this design incurs a massive spatial overhead, necessitating a forest of $2(n + m)$ parallel binary trees and long-range interactions to maintain time efficiency. Under nearest-neighbor constraints, the effective circuit depth becomes the same as that of BB, while requiring far more spatial resources.

Our model resolves this dichotomy: it matches the local-interaction time complexity and the passive structure of ASY but collapses the spatial footprint onto a *single* binary tree. By strictly enforcing local operations through our backup-walker protocol, we also eliminate the need for long-range connectivity.

Our approach: routing architecture

Encoding and scattering. In our qRAM model, the registers A and D of Eq. (1) are implemented via ordered sequences of quantum walkers, each possessing an internal “color” degree of freedom spanned by $|R\rangle$ (red) and $|B\rangle$ (blue), while the absence of a walker is denoted by $|\emptyset\rangle$. We emphasize that this empty state is to be understood as a spatial gap between colored walkers in the input sequence, rather than a distinct internal state. A bit $a_i = 0$ is encoded as $|\emptyset\rangle_{A_i}$, and $a_i = 1$ as $|R\rangle_{A_i}$. Since superpositions of addresses require superpositions of particle number, the simplest realization uses bosonic carriers; however, dual-rail and qudit-based encodings using fermionic or qubit carriers are also possible [16]. Each node of the binary tree implements a fixed scattering unitary $\hat{S}$: red walkers exit left (color unchanged), blue walkers exit right (color reset to red), and vacuum states propagate trivially (see Fig. 1a). This gate is address-independent and requires no external control signals.

Routing via $\hat{U}^{(i)}$. Before scattering at depth i , the i -th address walker transfers its routing information to all subsequent walkers via a gate $\hat{U}^{(i)}$: if A_i is red, it flips the color of all following walkers; if A_i is absent, nothing happens. After this conditional flip, $\hat{S}$ routes blue walkers right and red walkers

left, completing the address-dependent branching at level i . In its basic form, $\hat{U}^{(i)}$ is a multi-target controlled-NOT with a fan-out scaling linearly with the number of subsequent walkers. The only form of global coordination required is a classical clock that synchronizes the sequential application of $\hat{U}^{(i)}$ and $\hat{S}$ as walkers reach successive tree levels. This clock broadcasts no address information and does not constitute active, address-dependent routing.

The backup-walker decomposition

The implementation of $\hat{U}^{(i)}$ involves a many-body interaction whose range grows with the number of walkers, which is the main obstacle to a local implementation [18, 19]. Our central technical contribution is a *modular decomposition* of $\hat{U}^{(i)}$ into nearest-neighbor gates, at the cost of doubling the walker count. The initial state of the qRAM is prepared as in the standard protocol, but with a red backup walker following each primary address and memory walker. More precisely, after each walker W_i , we place its corresponding backup $\tilde{W}_i$ in the state $|R\rangle_{\tilde{W}_i}$. The decomposition at depth i proceeds in two steps, applied iteratively (see Fig. 1b):

1. **Initialization** ($\hat{U}_{\text{in}}$): A C-NOT between A_i and its backup $\tilde{A}_i$ transfers the address bit. If A_i is in the state $|R\rangle$, the backup flips to $|B\rangle$; otherwise, it remains $|R\rangle$.
2. **Propagation** ($\hat{U}_B$, repeated): A C-NOT-NOT gate acts on three consecutive walkers ($\tilde{W}_k, W_{k+1}, \tilde{W}_{k+1}$), activated when the control backup $\tilde{W}_k$ is blue. After each application, walkers shift forward by two positions. This is iterated $n + m - (i + 1)$ times at level i .

Both $\hat{U}_{\text{in}}$ and $\hat{U}_B$ are Clifford gates acting on at most three walkers with nearest-neighbor connectivity. One verifies by direct computation [16] that $\hat{U}_{\text{in}}$ followed by the iterated $\hat{U}_B$ reproduces $\hat{U}^{(i)}$ exactly on the primary walkers. This compositional structure is the core of the architecture: routing to arbitrary depth is realized by iterating a single local building block, with the tree depth entering only through the number of repetitions.

Conclusions and Outlook

We have presented a qRAM architecture that achieves optimal asymptotic scaling across all fundamental resources including circuit depth, spatial footprint, and active gate count. By relying on a **single binary tree** structure and **passive unitary dynamics**, our design overcomes the massive spatial bottlenecks of previous quantum-walker proposals without sacrificing their logarithmic speed advantage. Crucially, the routing protocol is constructed exclusively from **Clifford gates** involving interactions at a maximum distance of two, eliminating the need for complex active switching networks. The impact of these reductions is significant for fault-tolerant implementations. Recent independent analysis [20] highlights that, among the architectures analyzed, our protocol yields the “most optimal resource scaling for generating fault tolerant queries” when combined with distillation/teleportation techniques. While our model inherits the error scaling characteristics typical of bucket-brigade structures, the restriction to local Clifford operations significantly simplifies the error-correction overhead. We also note that the underlying routing mechanism is symmetric: running the protocol in reverse with appropriately prepared data registers enables coherent write operations, extending the applicability beyond quantum ROM to full qRAM. By bridging the gap between abstract query complexity and physical layout constraints, this work provides a concrete blueprint for the experimental realization of the first scalable quantum random access memory.

References

- [1] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15), October 2009. ISSN 1079-7114. doi: 10.1103/physrevlett.103.150502. URL <http://dx.doi.org/10.1103/PhysRevLett.103.150502>.

- [2] Guang Hao Low and Yuan Su. Quantum linear system algorithm with optimal queries to initial state preparation, 2024. URL <https://arxiv.org/abs/2410.18178>.
- [3] B. D. Clader, B. C. Jacobs, and C. R. Sprouse. Preconditioned quantum linear system algorithm. *Physical Review Letters*, 110(25), June 2013. ISSN 1079-7114. doi: 10.1103/physrevlett.110.250504. URL <http://dx.doi.org/10.1103/PhysRevLett.110.250504>.
- [4] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, January 2017. ISSN 1095-7111. doi: 10.1137/16m1087072. URL <http://dx.doi.org/10.1137/16M1087072>.
- [5] Ivan Kassal, Stephen P. Jordan, Peter J. Love, Masoud Mohseni, and Alán Aspuru-Guzik. Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proceedings of the National Academy of Sciences*, 105(48), 2008. ISSN 1091-6490. doi: 10.1073/pnas.0808245105. URL <http://dx.doi.org/10.1073/pnas.0808245105>.
- [6] James D. Whitfield, Jacob Biamonte, and Alán Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109(5):735–750, March 2011. ISSN 1362-3028. doi: 10.1080/00268976.2011.552441. URL <http://dx.doi.org/10.1080/00268976.2011.552441>.
- [7] Ryan Babbush, 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. *Phys. Rev. X*, 8:041015, Oct 2018. doi: 10.1103/PhysRevX.8.041015. URL <https://link.aps.org/doi/10.1103/PhysRevX.8.041015>.
- [8] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P.D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, October 2019. ISSN 0009-2665. doi: 10.1021/acs.chemrev.8b00803.
- [9] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, July 2014. ISSN 1745-2481. doi: 10.1038/nphys3029. URL <http://dx.doi.org/10.1038/nphys3029>.
- [10] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical Review Letters*, 113(13), September 2014. ISSN 1079-7114. doi: 10.1103/physrevlett.113.130503. URL <http://dx.doi.org/10.1103/PhysRevLett.113.130503>.
- [11] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum private queries. *Physical Review Letters*, 100:230502, Jun 2008. doi: 10.1103/PhysRevLett.100.230502. URL <https://link.aps.org/doi/10.1103/PhysRevLett.100.230502>.
- [12] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM Journal on Computing*, 35(1):170–188, 2005. doi: 10.1137/S0097539703436345. URL <https://doi.org/10.1137/S0097539703436345>.
- [13] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum random access memory. *Physical Review Letters*, 100(16), April 2008. ISSN 1079-7114. doi: 10.1103/physrevlett.100.160501. URL <http://dx.doi.org/10.1103/PhysRevLett.100.160501>.
- [14] Connor T. Hann, Gideon Lee, S.M. Girvin, and Liang Jiang. Resilience of quantum random access memory to generic noise. *PRX Quantum*, 2(2), April 2021. ISSN 2691-3399. doi: 10.1103/prxquantum.2.020311. URL <http://dx.doi.org/10.1103/PRXQuantum.2.020311>.
- [15] Ryo Asaka, Kazumitsu Sakai, and Ryoko Yahagi. Quantum random access memory via quantum walk. *Quantum Science and Technology*, 6(3):035004, May 2021. ISSN 2058-9565. doi: 10.1088/2058-9565/abf484. URL <http://dx.doi.org/10.1088/2058-9565/abf484>.

- [16] G. De Riso, G. Catalano, S. Lloyd, V. Giovannetti, and D. De Santis. A resource-efficient quantum-walker Quantum RAM, 2025. URL <https://arxiv.org/abs/2508.02855>. arXiv:2508.02855.
- [17] Samuel Jaques and Arthur G. Rattew. QRAM: A Survey and Critique. *Quantum*, 9:1922, December 2025. ISSN 2521-327X. doi: 10.22331/q-2025-12-02-1922. URL <https://doi.org/10.22331/q-2025-12-02-1922>.
- [18] Srinivasan Arunachalam, Vlad Gheorghiu, Tomas Jochym-O'Connor, Michele Mosca, and Priyaa Varshinee Srinivasan. On the robustness of bucket brigade quantum ram. *New Journal of Physics*, 17(12):123010, December 2015. ISSN 1367-2630. doi: 10.1088/1367-2630/17/12/123010. URL <http://dx.doi.org/10.1088/1367-2630/17/12/123010>.
- [19] Shifan Xu, Connor T. Hann, Ben Foxman, Steven M. Girvin, and Yongshan Ding. Systems architecture for quantum random access memory. In *56th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO '23*, page 526–538. ACM, October 2023. doi: 10.1145/3613424.3614270. URL <http://dx.doi.org/10.1145/3613424.3614270>.
- [20] Ansh Singal and Kaitlin N. Smith. Heterogeneously error-corrected qrams, 2025. URL <https://arxiv.org/abs/2504.21687>.