

Benchmarking a neutral-atom quantum computer

N. Wagner *,†, C. Poole†, T. M. Graham† and M. Saffman †,‡,§

**West High School, 30 Ash St. Madison, WI, 53726, USA*

†*Department of Physics,
University of Wisconsin-Madison, Madison, WI, USA*

‡*Inflection, Inc., Madison, WI, 53703, USA*

§*msaffman@wisc.edu*

Received 18 December 2023

Accepted 15 January 2024

Published 17 February 2024

In this study, we simulated the algorithmic performance of a small neutral atom quantum computer and compared its performance when operating with all-to-all versus nearest-neighbor connectivity. This comparison was made using a suite of algorithmic benchmarks developed by the Quantum Economic Development Consortium. Circuits were simulated with a noise model consistent with experimental data from [*Nature* **604**, 457 (2022)]. We find that all-to-all connectivity improves simulated circuit fidelity by 10%–15%, compared to nearest-neighbor connectivity.

Keywords: Quantum computing; benchmarking; neutral atom qubits.

1. Introduction

Many candidate platforms for the physical realization of quantum computing hardware have emerged in recent years, including neutral atoms, trapped ions, superconductors, quantum dots and photonics.^{1,2} All architectures suffer from various types of errors that functionally limit the number of qubits (circuit width) and gates (circuit depth) that can be executed in a circuit while maintaining high fidelity results.

Several different benchmarks have been proposed to compare the performance of quantum computers.^{3–6} We have chosen to use a benchmark developed by the Quantum Economic Development Consortium (QED-C benchmark), which uses the fidelities of several quantum circuits, such as the Quantum Fourier Transform and Grover’s algorithm, to quantify the performance of a computer.⁷

In this work, we present a noise model developed for a neutral atom quantum processor and benchmark simulated circuit execution. We constructed the

[§]Corresponding author.

noise model through a combination of diagnostic measurements and fitting simulated bitstring probabilities to measurements of circuits previously run on the computer. The quantum processor uses neutral atom qubits on a 2D square grid with gates implemented using lasers and microwave fields.⁸ Because some circuits in the QED-C benchmark require more qubits than were used in the quantum processor, we assumed the errors are uniform on each site and extended the simulator to accommodate high-width circuits provided in the benchmark. We ran simulations with both all-to-all and nearest-neighbor connectivity to accommodate the range of potential future device connectivity capabilities. We found significant improvement of circuit fidelities for the all-to-all connectivity compared to nearest-neighbor connectivity.

This paper is organized as follows. In Sec. 2, we discuss the methods by which we constructed the noise model for the simulation and give an overview of the QED-C benchmark. In Sec. 3, we present the results of the simulated benchmark circuits. We conclude in Sec. 4 with a discussion of the significance of the results.

2. Methods

2.1. Hardware

The neutral atom quantum processor we studied consists of a two-dimensional array of cesium atoms where qubits are encoded in the hyperfine clock states $|0\rangle = |6s_{1/2}, f = 3, m = 0\rangle, |1\rangle = |6s_{1/2}, f = 4, m = 0\rangle$. Single qubit R_z gates were applied with a local Stark shift provided by focused 459-nm light that was blue-detuned from the $|6s_{1/2}, f = 4\rangle \rightarrow |7p_{1/2}, f = 4\rangle$ transition. Controlled-Z (C_Z) gates were implemented using two-photon excitation of Rydberg states.⁹ Global rotations about an arbitrary axis in the $x - y$ plane of the Bloch sphere were performed with resonant microwave fields. The global rotations are written as $R_\phi^G(\theta)$, where ϕ is the angle between the rotation axis and the x -axis and θ is the rotation angle.⁸ These three gates provide a universal gate set from which any quantum circuit can be constructed. For example, local $R_\phi(\theta)$ rotations may be implemented by combining global microwave rotations and local Stark shifts using the identity

$$R_\phi(\theta) = R_{\phi+\frac{\pi}{2}}^G\left(\frac{\pi}{2}\right) R_z(\theta) R_{\phi+\frac{\pi}{2}}^G\left(-\frac{\pi}{2}\right). \quad (1)$$

This provides a rotation to the Stark shifted site, but the other atoms of the array see no net rotation since

$$I = R_{\phi+\frac{\pi}{2}}^G\left(\frac{\pi}{2}\right) R_{\phi+\frac{\pi}{2}}^G\left(-\frac{\pi}{2}\right). \quad (2)$$

The three gate types have different corresponding noise channels, as listed in Table 1. Independent experimental measurements determined T_1 and T_2 .⁸ We estimated the remaining noise parameters by fitting simulated bitstring populations against experimental populations of three 4-qubit QAOA-MaxCut circuits presented in the paper. The quality of the fit is determined by the classical fidelity

Table 1. List of the types of error used in the noise model, along with the probabilities of each being applied. The noise channels and rates are organized by gate. Note that in some cases multiple gates share the same noise channel at different rates. The errors for the R_ϕ^G and R_z gates are for π pulses. The indicated C_z gate fidelity in the last column is the entanglement fidelity obtained directly from the density matrix of a single application of the gate on a random pure state (not Hellinger distance), and is SPAM-corrected. For phase-dependent channels, the error/avg. gate fidelity is per π pulse. For additional details see Appendix A.

Gate/Process	Noise	Error	Average gate fidelity
Global R_ϕ^G	Depolarization	1.8×10^{-6}	0.9995
Local R_z	Phaseflip	3.2×10^{-4}	0.995
	Loss to Dark State	1.9×10^{-4}	
	Loss to Bright State	2.7×10^{-4}	
	Decay	2.0×10^{-8}	
	Phaseflip	3.3×10^{-2}	0.954
C_z	Loss to Dark State	1.8×10^{-2}	
	Loss to Bright State	2.9×10^{-2}	
	Decay	2.1×10^{-5}	
	Phaseshift	-2.0×10^{-3} rad.	
	Preparation	5.2×10^{-3}	NA
SPAM	Measurement	5.3×10^{-3}	
	$T_1 = 10$ s $T_2^* = 3.5$ ms $P_{ 0\rangle} = 0.42$ at $t = \infty$	NA	NA
Decoherence			

between the two probability distributions, defined as

$$F_s(P_{\text{ideal}}, P_{\text{out}}) = \left(\sum_x \sqrt{P_{\text{ideal}}(x)P_{\text{out}}(x)} \right)^2, \quad (3)$$

where x is summed over all possible bitstrings, P_{ideal} is the ideal output and P_{out} is the noisy output. This measure is then normalized to make the fidelity between the output state and the maximally mixed distribution 0. The resulting normalized fidelity is defined by

$$F_n(P_{\text{ideal}}, P_{\text{out}}) = \frac{F_s(P_{\text{ideal}}, P_{\text{out}}) - F_s(P_{\text{ideal}}, P_{\text{uni}})}{1 - F_s(P_{\text{ideal}}, P_{\text{uni}})}, \quad (4)$$

where the state P_{uni} is the maximally mixed multi-qubit state which has each diagonal element equal to $1/2^n$, with n the number of qubits, and all off-diagonal entries equal to zero. This normalized fidelity can be negative in certain cases, so the final reported fidelity is defined as

$$F(P_{\text{ideal}}, P_{\text{out}}) = \max(F_n(P_{\text{ideal}}, P_{\text{out}}), 0.0). \quad (5)$$

To quantify fidelities of the quantum gates F_G , we use the definition

$$F_G(\rho, \sigma) = [\text{Tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})]^2, \quad (6)$$

where ρ and σ are density matrices of the ideal and simulated states. This measure takes both phase and population errors into account. The average fidelity of a quantum gate is calculated by averaging fidelity between ideal and simulated outputs of a set of Haar-random input states. The gate fidelities are reduced by various sources of error present in the physical implementation of the quantum gates. All error channels are described further in Appendix A.

The two-qubit C_Z gate uses Rydberg blockade to entangle qubits by applying a phase shift conditioned on the state of the qubits. During this gate, both qubits are partially excited to a high-energy Rydberg state outside of the computational basis. An ideal C_Z operation maps the Rydberg state populations back to the computational basis at the end of the gate. In practice, the population transfer is imperfect, and some qubits are either lost from the trap or to states outside of the computational basis. For simplicity, we treat leakage and atom loss events as either populating a dark lost state $|l\rangle_0$ (for atoms lost from the array or leaked into $f = 3, m \neq 0$ states that appear dark during the state-selective readout), or a bright lost state $|l\rangle_1$ (for atoms leaked into $f = 4, m \neq 0$ states that appear bright during the state-selective readout), which does not participate in subsequent gate operations. To track the effects of these loss states in addition to the computational basis states, we simulate the computer as a ququart system rather than a qubit system. The C_Z gate is the most error-prone operation, with a simulated average gate fidelity of $\mathcal{F} \simeq 0.95$. The gate introduces a phaseflip channel, and a probability of qubit loss to a bright state $|l\rangle_1$ or dark state $|l\rangle_0$ from state $|1\rangle$. It also introduces a qubit decay channel, which transfers qubits in the $|1\rangle$ state to qubits in the $|0\rangle$ state.

The microwave and local R_z gates had the highest simulated fidelities with $\mathcal{F}_{R_\phi} \simeq 0.9995$ and $\mathcal{F}_{R_z} \simeq 0.995$. Microwave gate errors were modeled by following a perfect rotation with a probabilistic depolarization error. Local R_z gate errors were modeled by following a perfect rotation by a phase-flip, decay and loss channels to both dark and bright lost states. The T_1 decoherence time is 10^6 times longer than the average gate execution time, making it negligible. Both the $|0\rangle$ and $|1\rangle$ states can scatter into each other (primarily due to Raman scattering from the trap light), with a measured equilibrium population of $P_{|0\rangle} = 0.42$ in the $|0\rangle$ state. T_2 decoherence, which is ~ 100 times greater than gate execution times, is more significant, and significantly degrades the performance of longer circuits.

State preparation and measurement (SPAM) errors were simulated to occur with about 0.5% probability in both preparation and measurement. Preparation error acts as a global bit-flip channel proceeding the circuit, while measurement error randomly flips the bits of the output states before they are measured.

The average classical fidelity (see Eq. (5)) between experimental and simulated populations for the three circuits was 98.6%, with a minimum of 97.5% in the $p = 3$ MAXCUT circuit. The average Bell state fidelity with these parameters is 91.3%, which agrees with experimental results in Ref. 8. Note that in the circuit producing a Bell state in Ref. 8 a C_Z operation is performed on a state particularly sensitive to

noise introduced by the C_Z gate. On average, simulated GHZ circuits overestimated the fidelity of the GHZ state by 8.0%.

These discrepancies are likely due to the fact that noise parameters vary between sites and change over time in experiment, while we assume them to be constant in the model. Additionally, some noisy channels, particularly channels with relatively low error, were simplified to improve the simulator's overall complexity. In particular, the choice of depolarization for microwave rotations does not fully describe the behavior of the gate. Despite these simplifications, we see close correspondence between the simulated and measured circuit results.

2.2. Benchmark

We used the QED-C benchmark because it includes a variety of circuits and is thus an effective measure of the computer's practicality. Unlike quantum volume, which only uses square circuits,³ the QED-C benchmark uses both circuits of high depth and low width and circuits with low depth and high width. As such, we can independently test the limits of the processor with respect to depth and width. Other benchmarks, such as the benchmark utilizing mirror circuits described in Ref. 10, also sample more circuits at lower circuit widths. However, such benchmarks use impractical circuits that are randomly generated. It has been shown that the predictions given by randomized benchmarks, such as these, are not a good predictor of the fidelity of practical circuits.⁷

The QED-C benchmark uses the fidelity results of common quantum computing circuits to establish a means to compare quantum computers. The circuits used for benchmarking are divided into three broad categories: shallow oracle-based algorithms, quantum subroutines and circuits applicable to real-world problems. Each circuit is run one or more times at different widths and with random inputs. The fidelities from multiple runs are averaged and plotted based on the circuit width and the average transpiled circuit depth. Note that the QED-C benchmark uses two methods to benchmark the performance of the quantum Fourier Transform (QFT), which we will refer to as Method 1 and Method 2. Method 1 encodes a qubit into the Fourier basis using a QFT, modifies it, and decodes it with an inverse QFT. Method 2 prepares a qubit in the Fourier basis using a global Hadamard gate and R_z , then decodes it with an inverse QFT. In effect, Method 1 runs two consecutive QFTs, while Method 2 only runs one.

Since the computer did not have the capability to run some of the high-width circuits in Ref. 7, we sacrificed the accuracy of site-specific error parameters in the noise model for averages uniformly applied to all qubits. As shown in Fig. 1, the simulation still accurately described experimental results, with the additional benefit that it was able to simulate a larger quantum computer.

We simulated the computer's performance in two cases; a best case, where connectivity is all-to-all, and a worst case, where connectivity is limited to nearest neighbors on a two-dimensional square grid. The long range nature of the Rydberg

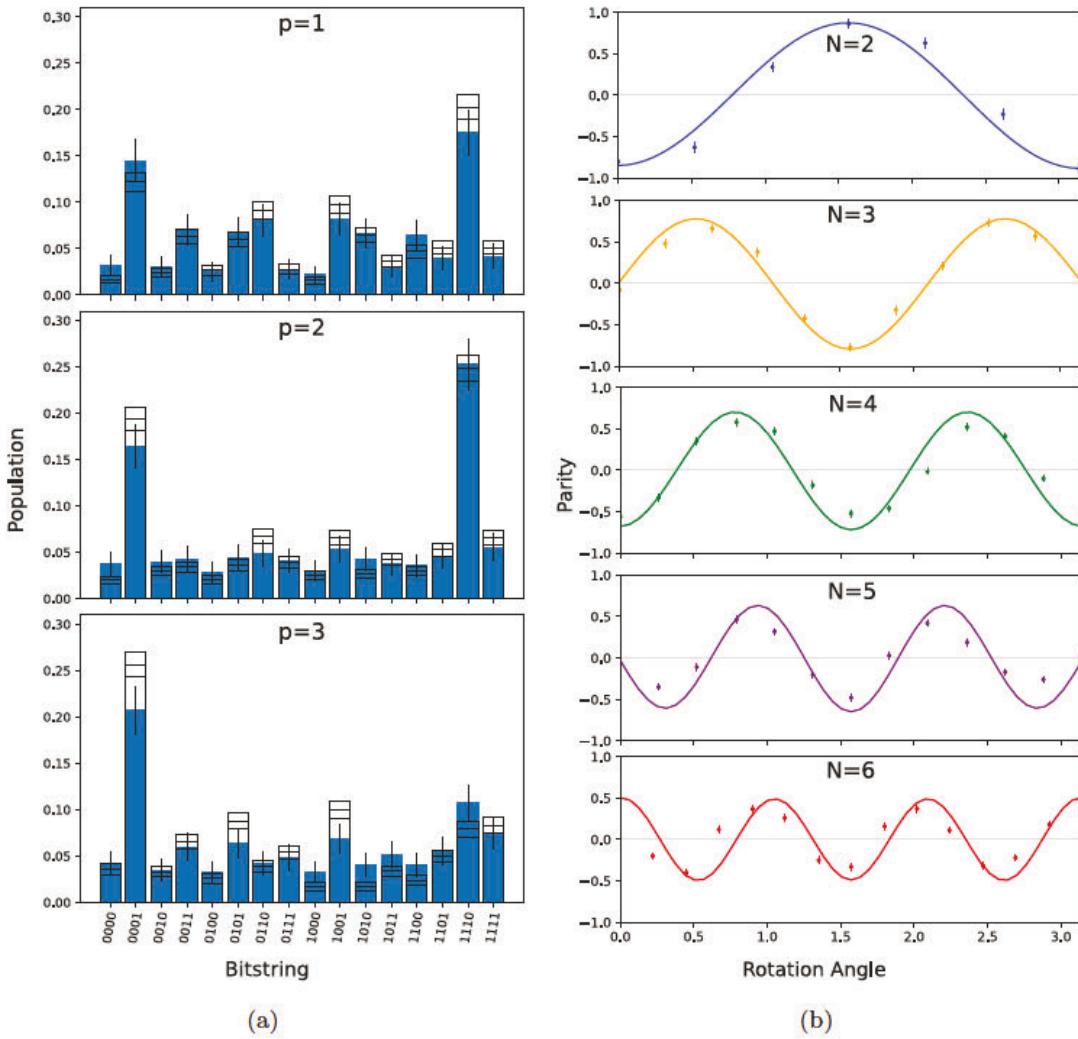


Fig. 1. (a) Simulation of 4-qubit MAXCUT algorithms from Ref. 8 with $p = 1$, $p = 2$ and $p = 3$. The unfilled bars represent experimental data and the blue fill represents simulated data. The lines represent error in the simulated data, and the dashed boxes represent error in the experimental data. (b) Parity oscillation curves of 2–6 qubit GHZ states. The data points represent experimental data and the solid-colored curves represent simulated data.

interaction makes it possible to consider beyond nearest-neighbor connectivity for moderate circuit widths. In this simulation, gate errors were assumed to be the same for both cases and independent of the distance between the two qubits. We simulated circuits with a width of up to 11 qubits.

3. Benchmark Results

As shown in Fig. 2, the computer simulation performed best with the shallower, oracle-based algorithms present in the QED-C benchmark (e.g. the Bernstein–Vazirani and Deutsch–Jozsa algorithms). With only nearest-neighbor connectivity, it ran the Bernstein–Vazirani and Deutsch–Jozsa circuits with up to 5 qubits with

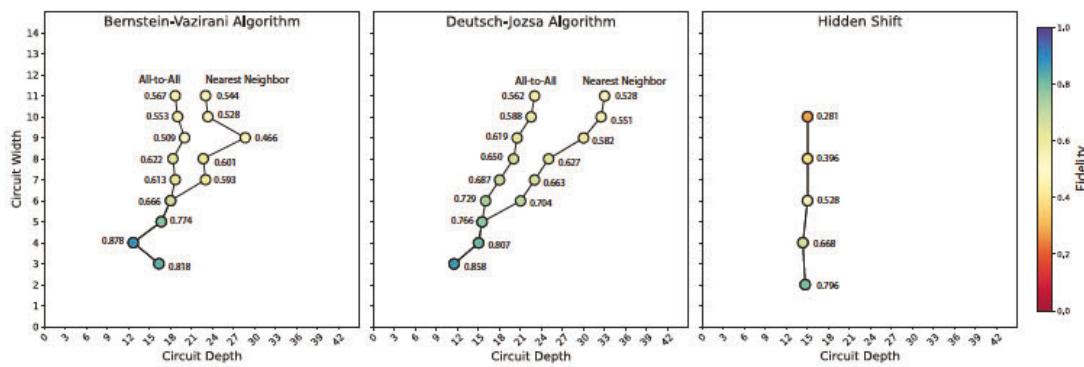


Fig. 2. Simulated results for the Deutsch–Jozsa, Bernstein–Vazirani, and Hidden Shift benchmarks for all-to-all and nearest-neighbor connectivity. For hidden shift there was negligible difference between all-to-all and nearest-neighbor circuits. The color bar shows the average fidelity of the result.

> 70% fidelity and the Hidden Shift circuit with up to 4 qubits with > 60% fidelity. Some accuracy was retained even with larger circuit widths, with a simulated 11-qubit implementation of the Bernstein–Vazirani and Deustch–Jozsa algorithms giving > 50% fidelity. Since the hidden-shift algorithm presented in the benchmark entangles only 2 qubits at a time, there is no difference between the performance of all-to-all versus nearest-neighbor connectivities in this circuit. However, all-to-all connectivity significantly increased the fidelity of wider circuits in the Deutsch–Jozsa and Bernstein–Vazirani algorithms. The most significant increases in fidelity occurred within the 7–11 qubit range, where the circuit fidelity increases on average by 2.6% (4.3% max) on the Deutsch–Jozsa circuits and 3.1% (3.7% max) on the Bernstein–Vazirani circuits.

The results of quantum subroutines are shown in Fig. 3. With only nearest-neighbor connectivity, Method 2 of the quantum Fourier transform benchmark reached a maximum of 85% fidelity with 2 qubits falling to 67% fidelity with 3 qubits. The simulation achieved an 89% fidelity with a 3-qubit phase estimation circuit. Method 1 of the QFT benchmark reached a maximum fidelity of 77% with 2 qubits falling to a 48% fidelity with 3 qubits. All amplitude estimation circuits produced results with fidelity < 20%. All-to-all connectivity also significantly improved the performance of quantum subroutines. The most significant increases in fidelity were found for Method 2 of the quantum Fourier transform benchmark and phase estimation. On average, the fidelity of the inverse quantum Fourier transform (Method 2) increased by 13% in the 3–5 qubit range, peaking at an increase of 17% at the 5-qubit implementation. The average increase in fidelity of phase estimation circuits on the same range was 22%, and peaks at a 26% increase at the 4 qubit implementation.

The results for several applications are shown in Fig. 4. Both the Grover's search and Hamiltonian estimation had high-fidelity results for 2 qubits, with both Grover's search and the Hamiltonian simulation giving 78% fidelity on average.

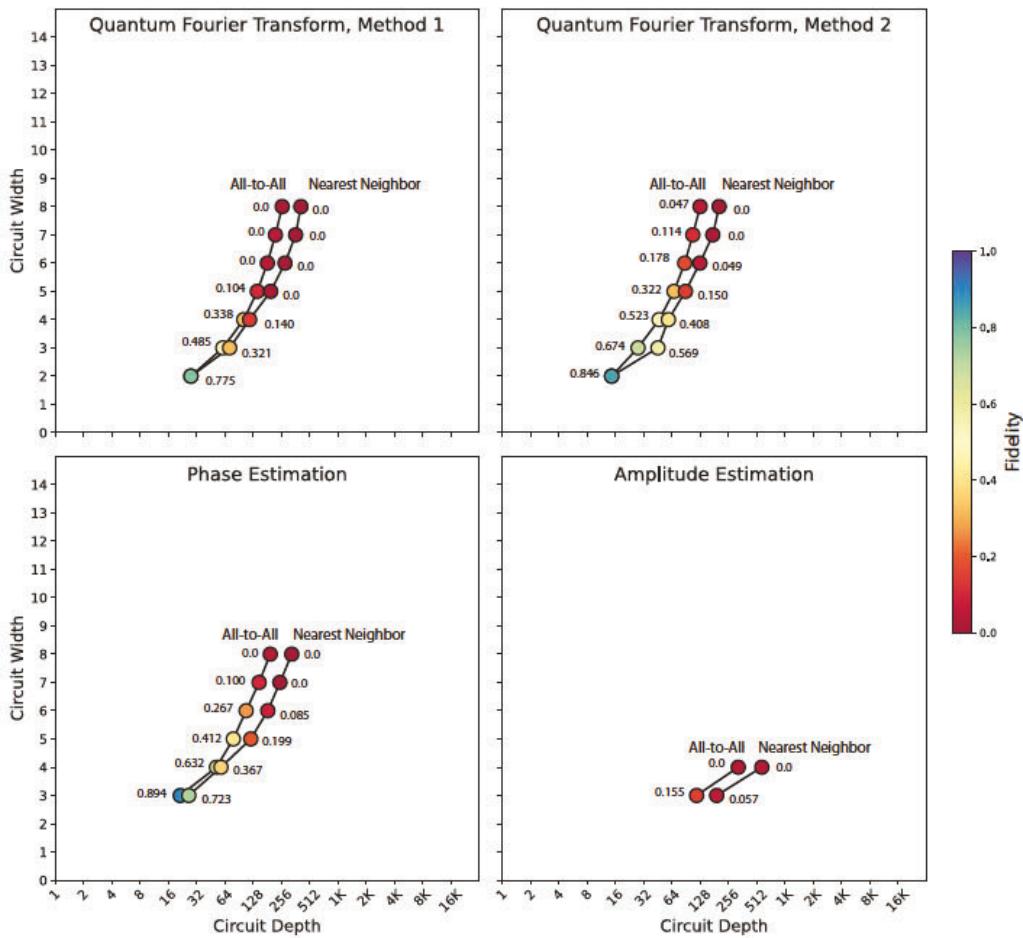


Fig. 3. Simulation results for the phase estimation, amplitude estimation, and both Quantum Fourier Transform benchmarks.

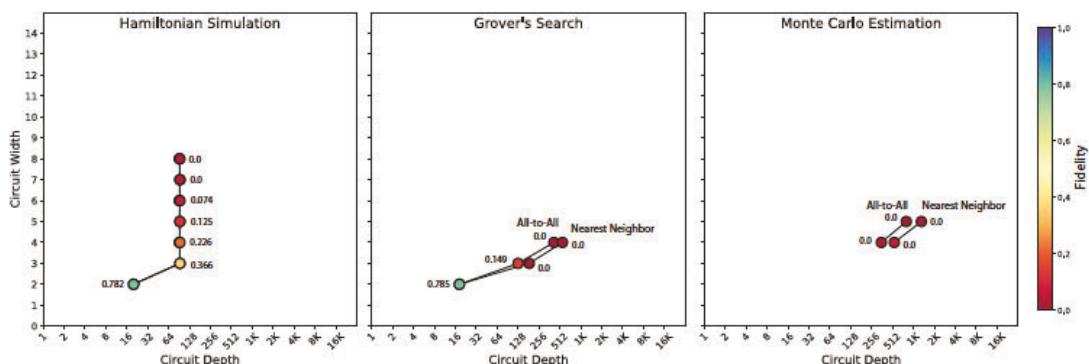


Fig. 4. Simulation results for Hamiltonian simulation, Grover's search, and Monte Carlo estimation. For Hamiltonian simulation there was negligible difference between all-to-all and nearest-neighbor circuits.

Implementations of the Hamiltonian simulation with 3–4 qubits are lower than that of the 2-qubit implementation, with the maximum fidelity being 36% in the 3-qubit implementation. All Hamiltonian circuits with 5 or more qubits, along with all other circuits with 3 or more qubits, returned an average fidelity of < 15% for both nearest-neighbor and all-to-all connectivity simulations. The nearest neighbor circuits were identical for all Hamiltonian circuits and the $n = 2$ Grover's search circuit, and all other results had too low a fidelity to be meaningfully compared between all-to-all and nearest neighbor implementations.

4. Conclusion

We have presented an error model which is consistent with experimental data and which can be used to simulate circuits with arbitrary widths and depths. When simulating the processor under different connectivity constraints, we found significant improvement in circuit fidelities as the connectivity of the computer increases. This improvement is consistent throughout several types of quantum circuits included in the QED-C benchmark, particularly with quantum subroutines and oracle-based algorithms.

Our inclusion of a proposed noise model provides an additional multifaceted, quantitative assessment of our computer's performance. It allows for us to determine specific sources of noise which are more prevalent than others, easily simulate how changes in one parameter may affect many different types of circuits, and instantly project the results of proposed experiments. In this way, using this noise model as a preliminary check in experimental work helps expedite the process of designing and improving our quantum processor.

Additionally, the benefits of a different proposed qubit topology than a square grid can be assessed without having to fundamentally change the computer. This could include studying and quantifying the benefits of a larger projected radius over which Rydberg blockage would occur or different layouts such as a triangular optical lattice. As developing a scalable quantum processor becomes more pertinent, the benefits and drawbacks of the qubit topology become more significant, making studies such as these more important to carry out.

We plan on continuing to study quantum subroutines on the computer, since they showed the most significant improvements. For example, the SupermarQ benchmark⁶ provides a set of mid-width circuits not used in the QED-C benchmark that would be interesting to run under our noise model.

Acknowledgments

This material is based upon work supported by NSF Award 2016136 for the QLCI center Hybrid Quantum Architectures and Networks, NSF award No. 2210437 and the U.S. Department of Energy Office of Science National Quantum Information Science Research Centers as part of the Q-NEXT center.

Appendix A. Noise Model

One can simulate a n -ququart lossy channel using a set of at most 16^n Kraus operators $\{\mathbf{A}_i\}$. The channel transforms the density matrix of the system as

$$\rho \rightarrow \sum_i \mathbf{A}_i \rho \mathbf{A}_i^\dagger, \quad (\text{A.1})$$

where $\sum_i \mathbf{A}_i^\dagger \mathbf{A}_i = \mathbf{I}$. The model uses 6 distinct channels in total, with both SPAM channels using the bit-flip error.

Throughout the appendix, we will generalize Pauli operators in the extended basis with a loss state $\{|0\rangle, |1\rangle, |l\rangle_0, |l\rangle_1\}$ as

$$\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (\text{A.2})$$

In general, the error probability p of a gate with an error rate r per π pulse, such as the R_ϕ and R_Z gates, is

$$p = r \frac{\theta}{\pi}. \quad (\text{A.3})$$

Parameters for error probabilities, and for T_1, T_2^* times, are listed in Table 1.

A.1. Depolarization error

A depolarization channel with error probability p acts upon a density matrix as

$$\rho \rightarrow (1-p)\rho + \frac{p}{3}\mathbf{X}\rho\mathbf{X} + \frac{p}{3}\mathbf{Y}\rho\mathbf{Y} + \frac{p}{3}\mathbf{Z}\rho\mathbf{Z}. \quad (\text{A.4})$$

This is the only type of error present on the microwave gates, and since the error rate is low, this type of error does not contribute significantly to the benchmark results.

A.2. Phase-flip error

A phase-flip channel with error probability p acts upon a density matrix as

$$\rho \rightarrow (1-p)\rho + p\mathbf{Z}\rho\mathbf{Z}. \quad (\text{A.5})$$

This noise channel is present on both the local R_z gate and the Rydberg gate.

A.3. Qubit loss error

The action of a qubit loss channel with error probability p on a density matrix can be represented as

$$\rho \rightarrow \mathbf{A}_0 \rho \mathbf{A}_0 + \mathbf{A}_1 \rho \mathbf{A}_1^\dagger, \quad (\text{A.6})$$

where

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (A.7)$$

and

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \sqrt{p} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (A.8)$$

for a loss to dark state channel and

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \sqrt{p} & 0 & 0 \end{bmatrix} \quad (A.9)$$

for a loss to bright state channel.

A.4. Qubit decay

Qubit decay decoheres qubits in the $|1\rangle$ basis to the $|0\rangle$ basis. A qubit decay channel with error rate p acts upon the density matrix as

$$\rho \rightarrow A_0 \rho A_0 + A_1 \rho A_1 \quad (A.10)$$

with

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & \sqrt{p} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (A.11)$$

A.5. Bit-flip

A bit-flip error operation is used to simulate qubit preparation error. The channel's action on the qubit can be given by

$$\rho \rightarrow (1-p)\rho + pX\rho X. \quad (A.12)$$

Qubit measurement error is performed after the bitstring populations are collected from the density matrix, reducing computation time. The error channel simply flips a bit in the measured bitstring with probability p for each bit.

A.6. Decoherence

Given a lossy density matrix

$$\rho = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix}, \quad (\text{A.13})$$

one can describe the action of the T_1, T_2^* decoherence channels directly on the matrix as

$$\rho \rightarrow \begin{bmatrix} d_1 \rho_{00} + (1 - d_1) p_0 p_t & d_2 \rho_{01} \\ d_2 \rho_{10} & d_1 \rho_{11} + (1 - d_1) p_1 p_t \end{bmatrix} \quad (\text{A.14})$$

with

$$d_1 = e^{-\frac{t}{T_1}}, \quad (\text{A.15})$$

$$d_2 = e^{-\frac{t}{T_2^*}}, \quad (\text{A.16})$$

$$p_1 = 1 - p_0 \quad (\text{A.17})$$

and

$$p_t = \rho_{00} + \rho_{11}, \quad (\text{A.18})$$

being the total population in the $|0\rangle$ and $|1\rangle$ states, where p_0 is the equilibrium population of $|0\rangle$ for the T_1 channel. To simplify the complexity of this channel for numerical computation, we describe it as a composition of two Kraus channels:

$$\begin{aligned} \rho &\rightarrow A\rho A^\dagger + B\rho B^\dagger + C\rho C^\dagger \\ &\rightarrow (1 - \phi)\rho + \phi Z\rho Z \end{aligned} \quad (\text{A.19})$$

with

$$A = \begin{bmatrix} \sqrt{p_0(1 - d_1) + d_1} & 0 & 0 & 0 \\ 0 & \sqrt{p_1(1 - d_1) + d_1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{A.20})$$

$$B = \begin{bmatrix} 0 & \sqrt{p_0(1 - d_1)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (\text{A.21})$$

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \sqrt{p_1(1-d_1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (A.22)$$

$$\phi = \frac{1}{2} - \frac{d_2}{2\sqrt{(p_0(1-d_1) + d_1)(p_1(1-d_1) + d_1)}}. \quad (A.23)$$

Note that, since the only mechanism for a qubit to enter the $|l\rangle_0$ or $|l\rangle_1$ state is the loss operator, all qubits in the $|l\rangle_0$ or $|l\rangle_1$ states will be incoherent with states where this qubit is in the $|0\rangle$ or $|1\rangle$ state. Therefore, instead of 16^n nonzero entries in the density matrix, there are only 6^n for each density matrix. This means that the circuits can be simulated much quicker than usual ququart implementations, since some channels require fewer Kraus operators.

The noise model we use always applies the gate itself first, and the T_1, T_2 channels last. For the microwave and local R_z gates, only one error is applied between the two operations. These errors are depolarization error and a phase-flip error, respectively. The C_Z gate applies the qubit loss channel directly after the gate, followed by the bit decay channel, the phase-flip channel, and finally the correlated phase-flip channel.

Appendix B. Methodology

Results were obtained by simulating the density matrix evolution of each circuit, with the Kraus operators described in Appendix A. Noise parameters were determined by maximizing the value of the average classical fidelities between the simulated and experimental results of the three QAOA circuits in Ref. 8, and are listed in Table 1. We chose the classical fidelity as our performance metric as it is directly connected to the metric used by the QED-C benchmark to test circuit accuracy.

The structure of the circuits provided in the benchmark depend on a variable parameter, meaning that two circuits of the same type and width can have different fidelities. Specific examples of this parameter include the phase to be measured in the phase estimation circuits, and the bitstring used in the Deutsch–Jozsa and Bernstein–Vazirani circuits. Because of this, for circuits that allowed 3 or more values for the parameter, we sampled 3 circuits per data point, corresponding to 3 distinct values chosen at random. We then took the resultant fidelity as listed in Figs. 2–4 as the average of the fidelities of the three runs. For circuits that admitted less than 3 values, we instead ran one circuit for each value. Circuits were sampled using the source code provided in Ref. 7. Exceptions to this triple sampling are the amplitude estimation circuits, where only two circuits per data point were chosen and Monte Carlo circuits, where only one circuit per data point was chosen to reduce processing time. These circuits had low fidelity in our simulations so further sampling was not required. Sample circuits are given in Figs. B.1 and B.2.

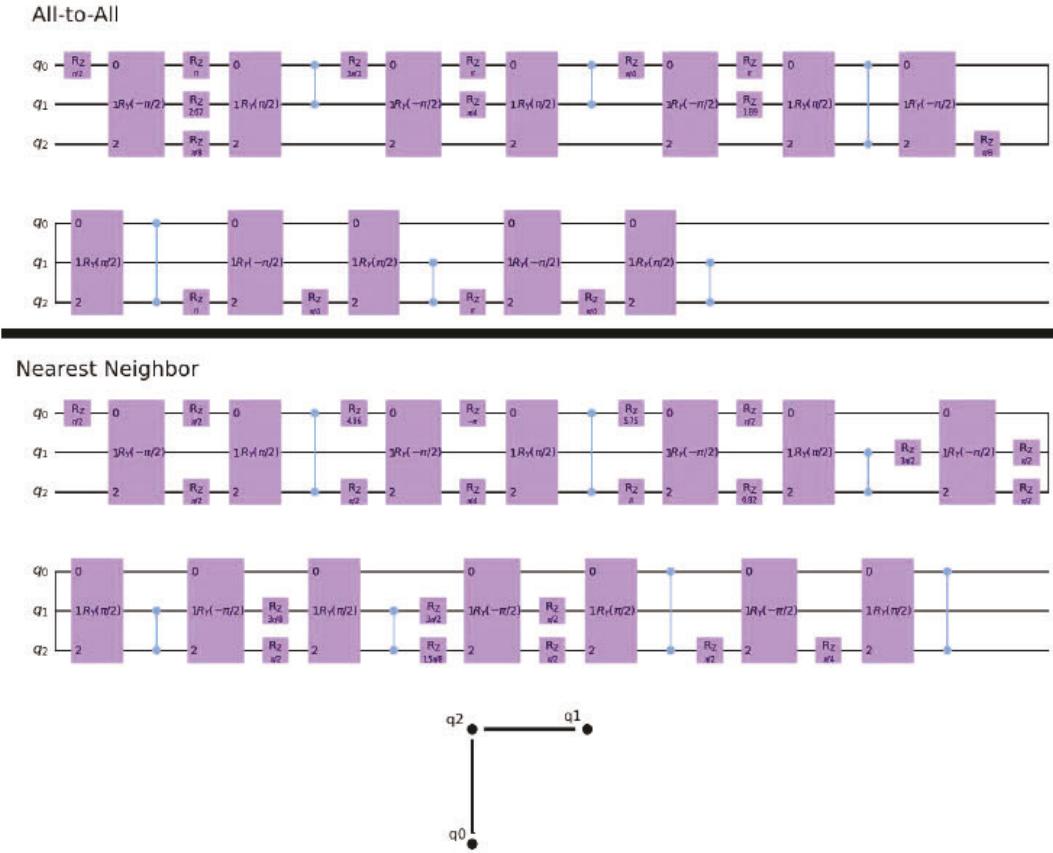


Fig. B.1. Sample circuit of a 3-qubit inverse quantum Fourier transform. Both the all-to-all and nearest-neighbor circuits are shown, with the nearest-neighbor topology shown below the circuit. The operations covering three qubits in one time slice are global microwave rotation gates $R_\phi^G(\theta)$.

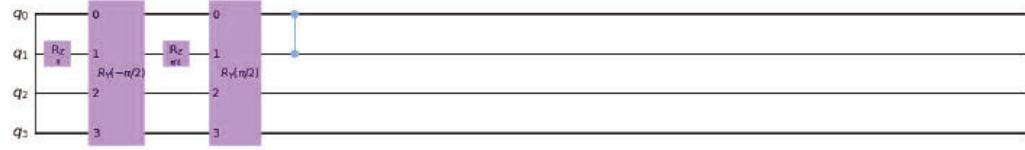
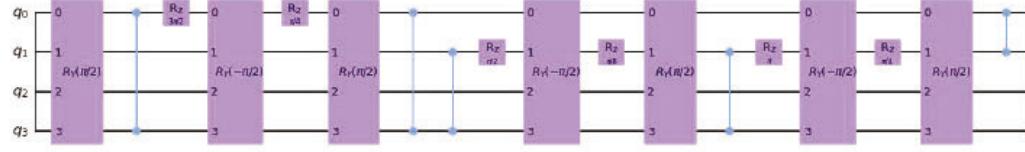
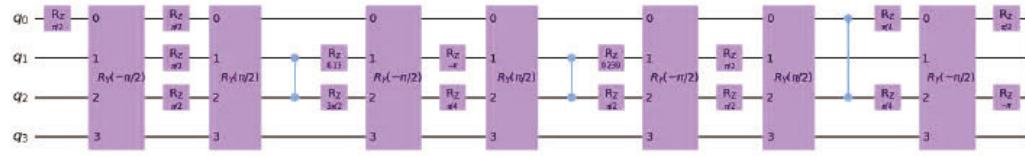
As the T_1 and T_2 times were measured precisely, both the T_1 and T_2 times were fixed, while the rest of the noise model parameters could vary. The goodness of fit between the two models was quantified by the classical fidelity between both distributions, which is closely related to the metric used in the benchmark. We optimized the parameters using the Nelder–Mead algorithm. As the simulator determines the exact resultant density matrix of the circuit, the populations of simulated states could be read directly after one evaluation of the circuit.

On average, the classical fidelity between the simulated and measured results of the three MAXCUT circuits is 98.7%, with the lowest being 97.5% on the $p = 3$ circuit. The fidelities of the simulated GHZ states overestimated the experimental fidelities by 8.0% and differed on average from the fidelities of the measured GHZ states by about 5.5%.

Appendix C. Treating Qubit Loss

The result of multiple runs of a benchmark circuit is a set of probabilities $P_{\text{out}}(x)$ for the observation of bit string x , with $\sum_x P_{\text{out}}(x) = 1$. Note that the fidelity used as a

All-to-All



Nearest Neighbor

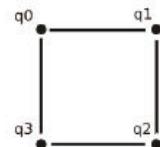
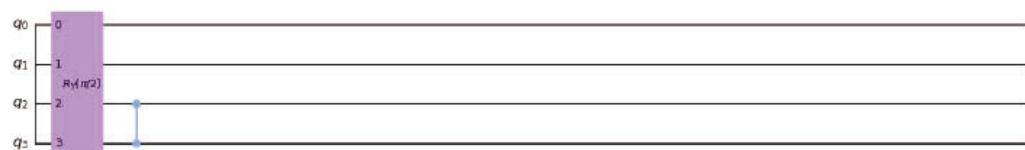
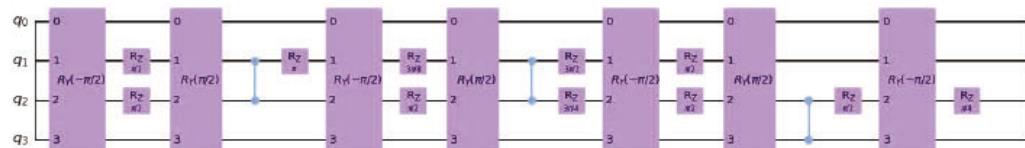
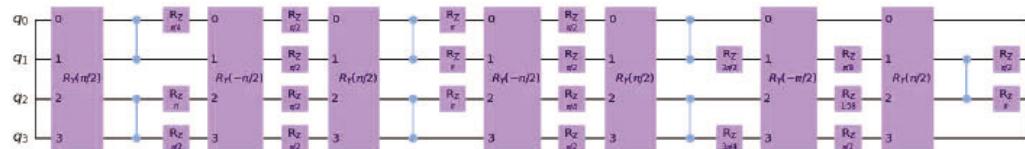
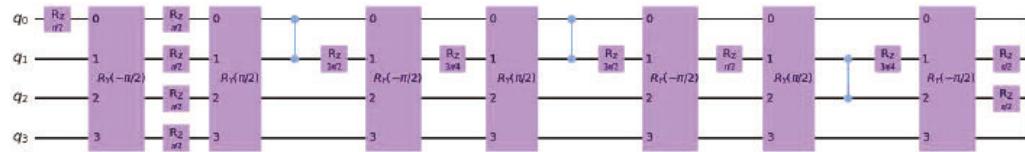


Fig. B.2. Sample circuit of a 4-qubit phase estimation circuit. Both the all-to-all and nearest-neighbor circuits are shown, with the nearest-neighbor topology shown below the circuit.

figure of merit in the QED-C benchmark only examines the probability distribution of reported bitstrings for a given circuit, not the final quantum state of the system. This is beneficial as it enables real hardware to be evaluated using the benchmark solely by repeatedly running circuits to acquire the probability distribution rather than using tomography or other similarly expensive means to reconstruct the final quantum state, but does introduce an insensitivity to certain types of errors. For example, any error in the relative phase of the computational basis states is masked completely.

Atoms that are lost from their traps during circuit execution will be interpreted by the atom readout procedure used in Ref. 8 as being in the $|1\rangle$ state. This is the fate of the majority of atoms captured by our catch-all bright lost state $|l\rangle_1$ and dark lost state $|l\rangle_0$. Thus, for both calibrating the error model and when calculating the fidelity of circuits in the benchmark, we reduce the ququart probability distribution to a qubit probability distribution by replacing the $|l\rangle_1$ state with the $|1\rangle$ state and the $|l\rangle_0$ state with the $|0\rangle$ state to mimic what we would receive as output from real hardware.

ORCID

Nathan Wagner  <https://orcid.org/0009-0001-8964-8975>

Mark Saffman  <https://orcid.org/0000-0001-6398-2097>

References

1. T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe and J. L. O'Brien, *Nature* **464** (2010) 45.
2. J. Bergou, M. Hillery and M. Saffman, *Quantum Information Processing: Theory and Implementation*, 2nd edn. (Springer, 2021).
3. A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation and J. M. Gambetta, *Phys. Rev. A* **100** (2019) 032328.
4. A. Wack, H. Paik, A. Javadi-Abhari, P. Jurcevic, I. Faro, J. M. Gambetta and B. R. Johnson, [arXiv:2110.14108](https://arxiv.org/abs/2110.14108).
5. S. Martiel, T. Ayral and C. Allouche, *IEEE Trans. Q. Eng.* **2** (2021) 3102011.
6. T. Tomesh, P. Gokhale, V. Omole, G. S. Ravi, K. N. Smith, J. Viszlai, X.-C. Wu, N. Hardavellas, M. R. Martonosi and F. T. Chong, SupermarQ: A scalable quantum benchmark suite, in *2022 IEEE Int. Symp. High-Performance Computer Architecture (HPCA)* (IEEE, 2022), pp. 587–603.
7. T. Lubinski, S. Johri, P. Varosy, J. Coleman, L. Zhao, J. Necaise, C. H. Baldwin, K. Mayer and T. Proctor, [arXiv:2110.03137](https://arxiv.org/abs/2110.03137).
8. T. M. Graham, Y. Song, J. Scott, C. Poole, L. Phuttitarn, K. Jooya, P. Eichler, X. Jiang, A. Marra, B. Grinkemeyer, M. Kwon, M. Ebert, J. Cherek, M. T. Lichtenman, M. Gillette, J. Gilbert, D. Bowman, T. Ballance, C. Campbell, E. D. Dahl, O. Crawford, N. S. Blunt, B. Rogers, T. Noel and M. Saffman, *Nature* **604** (2022) 457.
9. H. Levine, A. Keesling, G. Semeghini, A. Omran, T. T. Wang, S. Ebadi, H. Bernien, M. Greiner, V. Vuletić, H. Pichler and M. D. Lukin, *Phys. Rev. Lett.* **123** (2019) 170503.

10. K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Pisenti, M. Chmielewski, C. Collins, K. M. Hudek, J. Mizrahi, J. D. Wong-Campos, S. Allen, J. Apisdorf, P. Solomon, M. Williams, A. M. Ducore, A. Blinov, S. M. Kreikemeier, V. Chaplin, M. Keesan, C. Monroe and J. Kim, *Nat. Commun.* **10** (2019) 5464.