

Repeated Ancilla Reuse for Logical Computation on a Neutral Atom Quantum Computer

J. A. Muniz¹, D. Crow¹, H. Kim¹, J. M. Kindem¹, W. B. Cairncross¹, A. Ryou¹, T. C. Bohdanowicz¹, C.-A. Chen¹, Y. Ji¹, A. M. W. Jones¹, E. Megidish¹, C. Nishiguchi¹, M. Urbanek¹, L. Wadleigh¹, T. Wilkason¹, D. Aasen², K. Barnes¹, J. M. Bello-Rivas², I. Bloomfield¹, G. Booth¹, A. Brown¹, M. O. Brown¹, K. Cassella¹, G. Cowan¹, J. Epstein¹, M. Feldkamp¹, C. Griger¹, Y. Hassan¹, A. Heinz¹, E. Halperin¹, T. Hofler¹, F. Hummel¹, M. Jaffe¹, E. Kapit^{1,3}, K. Kotru¹, J. Lauigan¹, J. Marjanovic¹, M. Meredith¹, M. McDonald¹, R. Morshead¹, S. Narayanaswami¹, K. A. Pawlak¹, K. L. Pudenz¹, D. Rodríguez Pérez¹, P. Sabharwal¹, J. Simon⁴, A. Smull¹, M. Sorensen¹, D. T. Stack¹, M. Stone¹, L. Taneja¹, R. J. M. van de Veerdonk¹, Z. Vendeiro¹, R. T. Weverka¹, K. White¹, T.-Y. Wu¹, X. Xie¹, E. Zalus-Geller¹, X. Zhang¹, J. King¹, B. J. Bloom¹ and M. A. Norcia¹

¹*Atom Computing, Inc, Boulder, Colorado, USA*

²*Microsoft Quantum, Redmond, Washington, USA*

³*Department of Physics, Colorado School of Mines, Golden, Colorado, USA*

⁴*Department of Physics and Department of Applied Physics, Stanford University, Stanford, California, USA*



(Received 19 June 2025; accepted 16 October 2025; published 4 December 2025)

Quantum processors based on neutral atoms trapped in arrays of optical tweezers have appealing properties, including relatively easy qubit number scaling and the ability to engineer arbitrary gate connectivity with atom movement. However, these platforms are inherently prone to atom loss, and the ability to replace lost atoms during a quantum computation is an important but previously elusive capability. Here, we demonstrate the ability to measure and reinitialize, and if necessary replace, a subset of atoms while maintaining coherence in other atoms. This allows us to perform logical circuits that include single- and two-qubit gates as well as repeated midcircuit measurement while compensating for atom loss. We highlight this capability by performing up to 41 rounds of syndrome extraction in a repetition code, and combine midcircuit measurement and atom replacement with real-time conditional branching to demonstrate heralded state preparation of a logically encoded Bell state. Finally, we demonstrate the ability to replenish atoms in a tweezer array from an atomic beam while maintaining coherence of existing atoms—a key step toward execution of logical computations that last longer than the lifetime of an atom in the system.

DOI: [10.1103/v7ny-fg31](https://doi.org/10.1103/v7ny-fg31)

Subject Areas: Atomic and Molecular Physics,
Quantum Information

I. INTRODUCTION

Many prominent protocols for error-corrected quantum computing rely on the ability to measure the state of a subset of qubits—“ancilla” qubits—while maintaining quantum coherence in other “data” qubits [1,2]. This capability is referred to as midcircuit measurement (MCM). Motivated by this need, several demonstrations of MCM have recently been performed in systems composed of individually controlled neutral atoms trapped within optical tweezers [3–9]. In order to perform fault-tolerant circuits with multiple rounds of MCM, it is advantageous to be able to reuse ancilla qubits after reading out their state, requiring MCM to be both nondestructive and accompanied by reset of the

qubit state and motional state of the measured atoms. State readout of neutral atoms—including some previous demonstrations of MCM [5,8]—is often destructive, relying on the removal of one qubit state to obtain state selectivity. Other MCM protocols may be fundamentally compatible with ancilla reuse [4,7,9], but ancilla reuse within a quantum circuit involving a universal set of gate capabilities has not previously been demonstrated. This work demonstrates quantum circuits that contain repeated MCM with reset and reuse of ancilla qubits.

Even MCM protocols that are nominally nondestructive will suffer from occasional loss of ancilla qubits, either from the MCM process itself or from idle or operational losses in other parts of the circuit. To compensate for this loss, it is advantageous to be able to replace missing atoms after MCM. Multiple rounds of MCM can also be performed by replacing lost ancilla atoms in destructive MCM protocols, [9], though the much higher required rate of replacement presents challenges for large numbers of atoms or rounds of MCM. Here, we replace ancilla atoms that are

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

occasionally lost with atoms from a reservoir region. Because reservoir-based approaches will eventually run out of atoms, very long circuits also require a way to reload the reservoir while maintaining coherence in data qubits. We demonstrate this capability by refilling the reservoir from a spatially separated magneto-optical trap (MOT) (which itself is reloaded from an atomic beam), while maintaining coherence in qubits already within the system.

We showcase the ability to perform repeated reuse of measured atoms by performing multiple cycles of a repetition code, where ancilla atoms are reused between rounds of execution. This code is configured to swap the role of data and ancilla qubits on each cycle, allowing us to perform a delayed erasure conversion and correct for occasional loss of ancilla atoms through replacement from a reservoir while preserving logical states [10–12]. We perform up to 41 cycles of syndrome extraction while replacing ancillae as they are lost, observing error detection rates that are roughly constant versus the repetition index.

By combining MCM with ancilla reuse and classical branching, we demonstrate the ability to perform repeated attempts at fault-tolerant state preparation of logical Bell states, where the repetitions are performed until measurements of parity-check qubits indicate successful preparation of the desired state. Such capabilities form the basis of scalable state preparation, as multiple blocks can be built up one at a time, and stored as further blocks are prepared.

II. ZONED ARCHITECTURE AND MIDCIRCUIT MEASUREMENT

In this work, we demonstrate a zone-based quantum processor [8,10] that is capable of performing universal quantum computation as well as midcircuit measurement with qubit reset and reloading. This is facilitated by a combination of trapping wavelengths used in spatially separated zones to enable the local application of key operations. Our system uses nuclear-spin qubits formed within the $^1S_0$ ground states ($|0\rangle \equiv |F = 1/2, m_F = -1/2\rangle$ and $|1\rangle \equiv |F = 1/2, m_F = +1/2\rangle$) of ^{171}Yb atoms [10]. Quantum gates and atom movement are performed using optical tweezer arrays, while high-fidelity, nondestructive state-selective readout is performed by transferring atoms to collocated cavity-enhanced optical lattices [13]. Imaging is performed by collecting photons scattered from the relatively narrow (180-kHz linewidth) $|1\rangle$ to $^3P_1$ $F = 3/2$, $m_F = 3/2$ transition at 556 nm on a sensitive camera. Importantly, our imaging technique [6,14] enables unambiguous determination of both the qubit state and atom presence. Further, key gate errors are converted to atom loss [10]. The combination of these features allows us to perform (delayed) erasure conversion [11,12,15].

The zones and their functions are illustrated in Fig. 1(a), and key atomic states and gate lasers are shown in Fig. 1(b). The *register* is a region of 128 optical tweezers arranged in eight pairs of adjacent columns (16 sites per row), separated

by gaps that facilitate atom movement. The register holds atoms during a quantum circuit, and arbitrary site-selective single-qubit (1Q) gates are applied in the register with a pair of Raman beams [8,10,16,17]. The 423-nm wavelength of the register tweezers induces substantial light shifts on the imaging transition, which suppresses scattering when atoms in other zones are imaged (see Appendixes B and C for details). This is possible because both register tweezers and cavity lattices are present during midcircuit measurements.

A second group of 80 tweezers operate at 459 nm and have a matching column distribution to the register. These tweezers serve multiple functions: A single row forms the *interaction zone* (IZ), trapping atoms during our two-qubit (2Q) controlled- z (CZ) entangling gates via sequential excitation to a Rydberg state [16]; two rows make up the *measurement zone* (MZ), where atoms are placed to be measured during MCM; and the final two rows function as the *storage zone* (SZ), which acts as a nearby reservoir to replenish any missing atoms in the MZ after MCM. The 459-nm tweezers provide equal light shifts to the $^1S_0$ ground and $^3P_0$ states [18], which is key for executing the 2Q gates [16].

Finally, the *loading zone* (LZ) consists of 75 sites at 483 nm [6]. This densely packed array’s main purpose is to load and prepare fresh single atoms that are transferred from a magneto-optical trap situated 300 mm below the tweezer arrays, as detailed in Ref. [13].

Arbitrary gate connectivity across the system is enabled by atom movement [8,10]. These movements are executed using a set of 488-nm tweezers, controlled by a pair of crossed acousto-optical deflectors, utilizing the empty “highways” situated between column pairs within these arrays to move atoms while minimizing disturbances on qubits in the register. All atom movements occur while the cavity lattice is off.

MCM in our zoned architecture relies on the ability to image the atoms in the MZ and SZ with high fidelity, perform cooling and state preparation for atoms in the MZ and SZ after imaging, and move atoms from the SZ to the MZ to correct for detected loss, while minimizing loss and retaining the coherence of superposition states within atoms in the register. Imaging is performed using a global beam that illuminates all atoms (due to geometrical constraints). Cooling after imaging, as well as optical pumping, are implemented using local beams that illuminate only the IZ, SZ, and MZ. The large (≈ 100 – MHz) shifts induced by the register tweezers dramatically reduce scattering among register atoms from the global imaging beam and stray light from the local beams.

A typical *MCM cycle* is described in Fig. 1(a). It begins by moving subsets of ancilla qubits from the register to the MZ tweezers. Atoms from the MZ and SZ are then transferred from the 459-nm tweezers to the lattice for imaging. A first image serves to identify atoms in $|1\rangle$. This is immediately followed by an *MCM reset* cycle

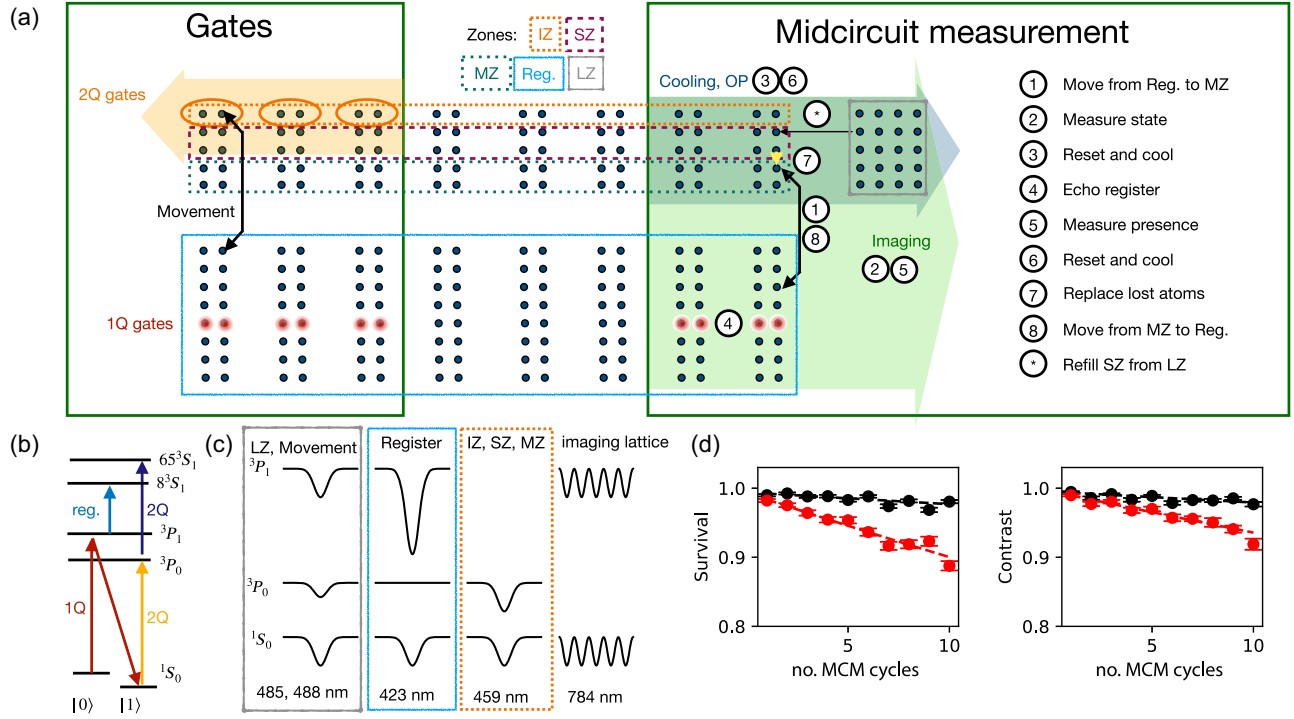


FIG. 1. (a) Zoned architecture for universal quantum computing. Optical tweezers are divided into subarrays, or zones, labeled interaction zone (IZ), storage zone (SZ), measurement zone (MZ), register (Reg.), and loading zone (LZ). Atoms can be transferred from tweezers to a collocated cavity-enhanced lattice (not shown) for readout. Two-qubit (2Q) controlled-z (CZ) gates are performed in parallel on all populated pairs of tweezers within the IZ (2Q gate lasers are represented by orange arrow). Site-resolved arbitrary single-qubit (1Q) gates are performed within the register (1Q gate lasers are represented by red circles). Arbitrary connectivity for 2Q gates is provided by moving selected atoms to desired positions within the IZ. Midcircuit measurement (MCM) allows state-selective, nondestructive imaging of the atoms within the IZ, SZ, and MZ, while maintaining coherence within the register. MCM is followed by cooling, qubit reset, and replacement of missing atoms with atoms from the SZ. The imaging laser is represented by the green arrow, while the lasers used for optical pumping and cooling are represented by the blue arrow. All beams shown address the full width of the array, but are depicted over a restricted portion for visual clarity. A full MCM cycle is depicted, and described in detail in the main text. (b) Selected energy levels of ^{171}Yb . Qubit states are defined by the nuclear spin of the ground state $1S_0$. Single-qubit gates are implemented using two-photon Raman transitions near detuned from the $3P_1$ manifold. Two-qubit gates are performed by sequential excitation from $|1\rangle$ to a high-lying Rydberg state, through a metastable $3P_0$ state [16]. Register traps are operated near resonance with the transition from $3P_1$ to $8s^3S_1$, leading to large shifts to $3P_1$ states that suppress unwanted scattering. (c) Qualitative polarizabilities of relevant states at wavelengths used for generating the different arrays, represented schematically as depths of trapping potentials. Within the LZ and movement tweezers, as well as the imaging lattices, $1S_0$ and $3P_1$ have equal polarizability. Within the register, $3P_1$ has large polarizability, while $3P_0$ is untrapped. In the tweezers that form the IZ, SZ, and MZ, $1S_0$ and $3P_0$ have equal polarizability. For visual clarity, the polarizability of states that are not addressed in a given trap are omitted. (d) Performance of MCM. Contrast and atom loss in the register are extracted from a Ramsey sequence containing a variable number of MCM cycles. Red markers correspond to the full MCM sequence. Black markers correspond to the same timing and operations (including lattice handoffs), but without imaging, cooling, or optical pumping light applied. The error bars represent the standard deviation of the mean, averaged over 128 measured qubit locations. In total, the MCM cycle contributes a fractional loss of 0.0098(7) and a contrast loss of 0.0062(5) among surviving atoms.

consisting of gray-molasses cooling [19] and optical pumping that prepare atoms in the MZ and SZ into state $|1\rangle$. A spin echo is performed on atoms in the register to cancel shifts from the MCM process [6,20] (see Appendixes B and C for details). A second image of state $|1\rangle$ confirms the presence of atoms regardless of their initial internal state, followed by a second reset to cool and prepare MZ and SZ atoms into state $|0\rangle$, which is not excited in our 2Q gates [16].

In Fig. 1(d) we show the typical atom survival and maintenance of coherence for atoms in the register during

MCM by inserting repeated MCM cycles in a Ramsey sequence. We observe atom loss of 0.0098(7) per cycle, and a site-averaged contrast loss of 0.0062(5) for atoms that survive in the register. Separately, we measure the survival of atoms in the MZ under MCM, and find a per-cycle loss of 0.005(2). We discuss the origin of this loss in Appendix C.

After the MCM block, the second MCM image is analyzed in real time by a software service that determines atom presence in the MZ and SZ, and programs the movement tweezers to fill detected vacancies within the

MZ with atoms from the SZ. We call such movements from filled sites to empty sites *conditional movements*. Further details can be found in Appendix B.

III. ATOM REUSE FOLLOWING MCM

In order to reuse atoms within a circuit following an MCM cycle, they must be reinitialized into a specific internal state (we use $|0\rangle$) and be sufficiently cold to enable high-fidelity gates. The imaging itself leaves the atoms slightly too warm for this, so we apply gray-molasses cooling after imaging to reduce the radial motional quantum numbers to $n \simeq 0.5$, which is compatible with high-fidelity gates in our system [10,16]. Here, we perform two gate performance benchmarks that compare atoms prepared in our standard register imaging and state preparation protocol “register preparation” (non-MCM as in Ref. [10]) versus cases where the atoms are prepared using an MCM cycle “MCM preparation.” In both cases, measurement and state preparation are performed in the optical lattice. The key differences between the methods are the location within the lattice where the operations are performed, the specific laser beams used (global vs local), and whether the register tweezers are turned on.

First, we perform a measurement that is sensitive to both 1Q and 2Q gate fidelities using a variation of the global-echo randomized benchmarking (GERB) protocol. In this protocol, concatenated circuit blocks each contain a random 1Q operation and two CZ gates separated by an echo pulse, as depicted in Fig. 2(a) and detailed in Refs. [10,16]. This method is similar to those used to characterize 2Q gate performance in other cold-atom systems [21]. As in Ref. [10], 1Q gates are performed in the register while 2Q gates are performed in the IZ. A final 1Q operation brings pairs of atoms to the $|11\rangle$ state in the absence of errors.

By observing the decay of population returned to $|11\rangle$ (and also of atom-pair survival) versus the number of blocks, we extract the combined error rate for the operations in the block [Fig. 2(b)]. We compare these error rates for register preparation and MCM preparation in Fig. 2(c), and find that MCM state preparation does not degrade performance. Total error per GERB block is 0.030(3) and 0.032(3) for MCM and register state preparation methods, respectively. When postselecting on atom survival, the error per GERB block is 0.012(1) and 0.016(1) for MCM and register state preparation methods, respectively. The typical total error probability per block for the same GERB sequence, but with the CZ gates omitted, is 0.008(2).

Our second benchmark characterizes our ability to generate the Bell state $|\Phi_+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ after each of the two preparation methods using modified versions of the circuits in Ref. [22]. In the MCM case, we prepare the atoms in the MZ, then move them to the register, and then apply the same Bell state generation circuit as when the atoms are prepared directly in the register. The fidelities

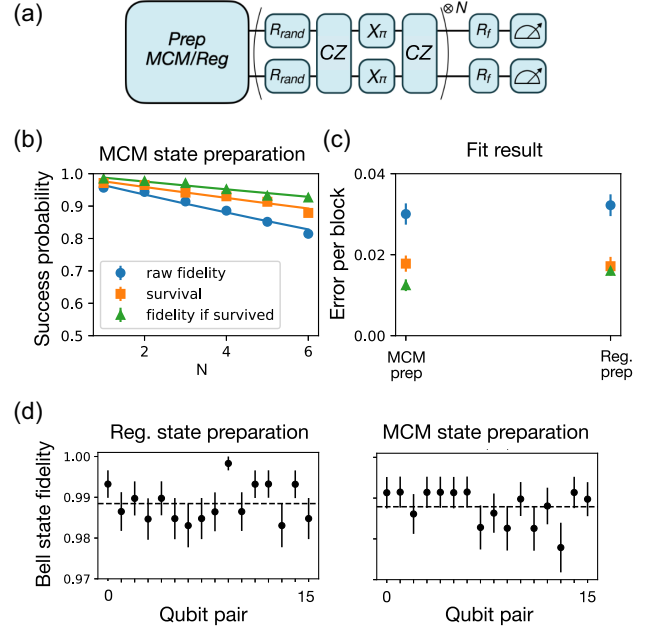


FIG. 2. Gate performance following MCM reinitialization. (a) The global-echo randomized benchmarking (GERB) sequence consists of N repeated blocks of 2Q and random 1Q gates performed on pairs of qubits, followed by a precomputed operation to bring population back to the $|11\rangle$ state in the absence of errors. We perform this sequence using atoms prepared using an MCM cycle in the MZ (MCM prep), or with atoms prepared directly in the register (Reg. prep). (b) Success probability versus N for MCM prep. By measuring both qubit states and survival [10,16], we separately extract total success probability (blue), cases where both atoms survive (orange), and the probability of measuring the correct two-qubit state when both atoms survive (green). The error bars represent the standard deviation of the mean, averaged over different random circuits. (c) Error rate per block, as extracted from exponential fits in (b) (and similar data for register prep), indicating compatible performance between the two preparation methods is shown on the right. The error bars represent the standard deviation of the fitted error rate. (d) Bell state fidelities for atoms initialized using register preparation (left) and MCM state preparation (right), plotted versus qubit pair indices for qubits within the top two rows of the register. The error bars represent the standard deviation calculated from a binomial distribution on the results from three measured bases, propagated to the reported fidelity.

(as defined in Ref. [23]) for both cases are shown in Fig. 2(d). We observe an average Bell state fidelity of 98.8(1)% for both preparations. These Bell state fidelities are postselected on atom survival [10], and so do not include additional loss associated with MCM cycle itself, which is characterized above. Additional loss in CZ gates for the MCM-prepared case is bounded by the GERB benchmark data of Fig. 2(b). The site-averaged probability that at least one atom in a Bell pair was identified as not present during the final readout is 0.017 for register preparation and 0.03 for MCM preparation. This difference

can likely be attributed to the additional imaging loss present in MCM imaging.

IV. REPETITION CODE WITH ATOM REUSE AND REPLACEMENT

We demonstrate the integration of MCM and atom reuse with a circuit-based error-correction protocol. We encode a single classical logical bit in a repetition code using N physical qubits (code words $|0\rangle^{\otimes N}$ and $|1\rangle^{\otimes N}$). We call this the distance of the code. We use N additional physical qubits to simultaneously perform syndrome extraction and loss detection through swap-based syndrome extraction. This circuit is constructed by adding SWAP gates to a standard syndrome extraction circuit and then compiling it into our native gate set [10,11]. The final circuit has the same number of 2Q gates as a standard syndrome extraction circuit. Although the standard repetition code requires only $N - 1$ parity checks, we implement a “ring code” or 1D toric code by adding one additional check. This additional check facilitates constructing a circuit in which a

physical qubit is active for at most two cycles before being measured, reinitialized, and, if necessary, replaced. For further information and raw data used to generate plots, see Appendix E.

We focus here on a bit-flip variant of the repetition code, though in our native gate set a phase-flip variant differs only by the insertion of 1Q gates that resolve to identity. Typically, this type of repetition code places each qubit in a computational basis state during MCM and so is insensitive to phase errors induced by MCM on data qubits. To make the code sensitive to such phase errors (which must be small for a quantum error-correcting code), we also characterize a second variant of the code that has a pair of Hadamard gates added to data qubits surrounding each MCM, placing the data qubits in a superposition state during MCM. We refer to these two variants of the code as *phase-insensitive* and *phase-sensitive*, respectively [Fig. 3(a)].

We performed repetition code memory experiments for various distances and numbers of cycles. In particular, we scanned odd distances between 3 and 9 while keeping the number of rounds equal to the distance. For these

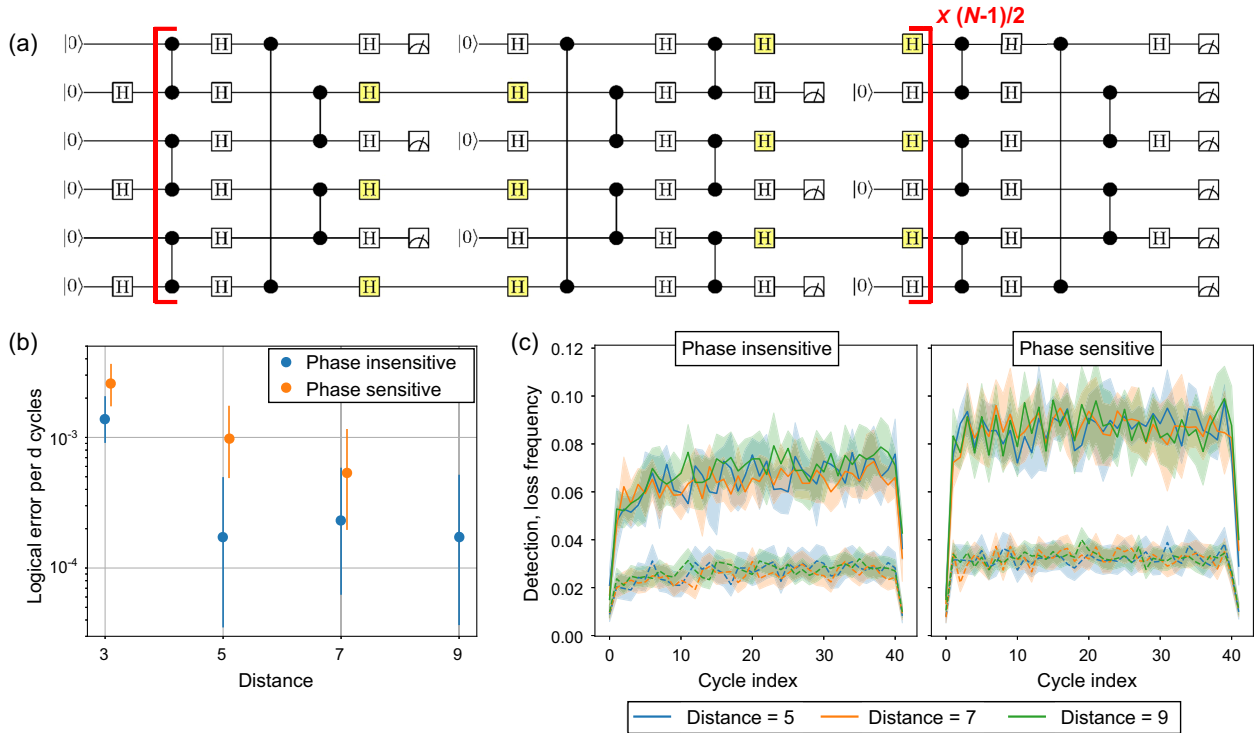


FIG. 3. (a) Distance 3, N cycle, phase-sensitive repetition code in the H and CZ basis. Information is encoded in the physical state of N qubits which undergo cyclic parity checks with N ancilla qubits. Physically, this circuit is compiled in the machine native gate set: $[R_z(\phi), R_x(\pi/2), CZ]$ and an additional X gate is applied to the nonmeasured qubits to echo noise during MCM. The highlighted H gates are removed for the phase-insensitive case. The portion of the circuit within the red brackets is repeated $(N - 1)/2$ times. (b) Demonstration of error suppression as a function of code distance computed for experiments with number of cycles equal to distance. Each data point consists of at least 11000 shots. The error bars show 95% confidence intervals. Data used to generate this plot can be found in Appendix E, Table IV. (c) Detection frequency (solid lines) and average per-qubit loss (dashed lines), along with their standard deviations (shaded areas), are calculated from 600 shots of the repetition code with 41 iterations and depth 3, 5, and 7. For phase-sensitive circuits (left-hand panel) we can see slightly elevated detection relative to phase-insensitive circuits (right-hand panel), as expected from the additional 1Q gates and sensitivity to phase errors occurring during the MCM and reinitialization steps.

experiments, the results were decoded using minimum-weight perfect matching implemented using PyMatching [24]. The matching graph was modified on a per-shot basis to account for atom loss. For more information, see Appendix E. A plot of logical failure rate versus distance is shown in Fig. 3(c). Though we observe a general trend of improved logical error rate versus distance, we refrain here from making detailed claims about threshold behavior or saturation due to the limited statistical precision of our dataset.

In addition to tracking logical performance, we conducted separate memory experiments for odd distances between 5 and 9 while repeating the syndrome extraction procedure for 41 rounds. For these experiments, we compute the frequencies of detection events [20,25] in order to examine the system behavior. These “detectors” represent the overall parities between consecutive syndrome checks on corresponding pairs of data qubits. Unlike bare syndrome measurements, detectors localize regions of error sensitivity.

In Fig. 3(b), we plot the average detection frequency per cycle for the phase-insensitive and phase-sensitive versions of the repetition code, respectively. The detection frequencies in the bulk remain relatively flat across multiple rounds and are independent of the distance (within the error bars). Crucially, these observations indicate that by swapping the roles of data and ancilla qubits, and by recooling (and if needed, replacing) measured atoms, repeated rounds of midcircuit measurement do not degrade circuit performance. Further, the constant detector frequency versus code distance indicates negligible penalty associated with increasing the number of qubits involved in a circuit for the qubit numbers explored here—an important feature for a scalable platform.

V. HERALDED FAULT-TOLERANT STATE PREPARATION WITH CONDITIONAL REPETITION

Fault-tolerant preparation of high-quality resource states is an essential feature of practical quantum computation. Leveraging our MCM and ancilla reuse scheme, we demonstrate heralded preparation of logical Bell states. This is achieved by iteratively repeating a $[[2, 1, 2]]$ distillation protocol encoded into $[[4, 2, 2]]$ codes. We prepare logical Bell pairs $|\bar{\Phi}_+\rangle$ on a data block and two ancillary blocks—one to identify bit-flip errors and another to identify phase errors. We repeat the distillation protocol until a set of MCM measurements on the ancillary blocks indicates successful state preparation of a target logical state on the data block. The defining characteristic of this demonstration is the execution of a conditional loop that repeats until the success criterion is met. Decoding of the ancilla and data blocks is performed in real time, on every shot, and does not rely on postselection. By using MCM to measure the ancilla blocks, we preserve the logical state on

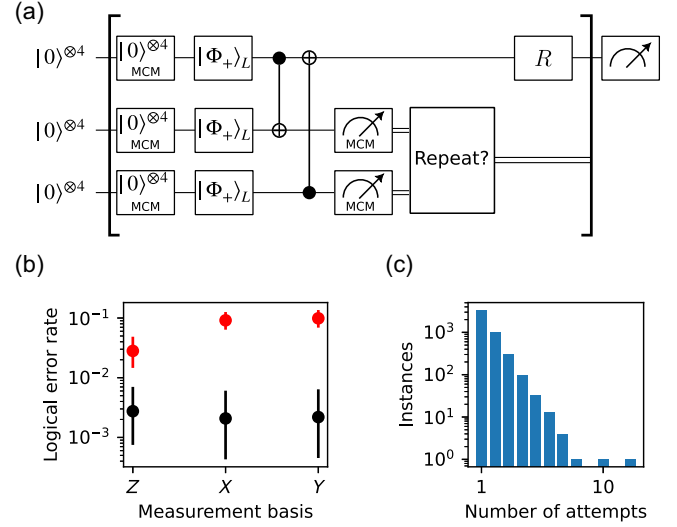


FIG. 4. (a) Fault-tolerant state preparation of logical Bell states encoded in the $[[4, 2, 2]]$ code through conditional retries. At the beginning of each state preparation attempt, data and ancilla qubits are reset via an MCM operation ($|0\rangle_{\text{MCM}}^{\otimes 4}$) and a set of logical Bell states $|\bar{\Phi}_+\rangle_L = |\Phi_+\rangle_L \otimes |\Phi_+\rangle_L$ are prepared in the data and ancilla blocks (four physical qubits each). After preparation ancillae are entangled with the data qubits and then measured using MCM. After a decoder assesses if the correct parity was measured, the data qubits are rotated to the desired basis using the rotation R . If the parity was incorrect, another attempt is performed. If the parity was correct, the retry loop exits and the logical state is measured. (b) Logical error rate on different measurement basis for instances where the decoder succeeded (black), and when it did not (red). For instances where the decoder did not succeed, the loop does not exit and we use the reset measurement of the following round to determine the logical error rate. Error bars represent the 95% confidence interval. Data used to create this plot can be found in Appendix F, Table V. (c) Histogram representing the number of attempts required to obtain a successful parity check. The average number is 1.44.

the data block, which remains available for use after measurement and classical logic are applied, as illustrated in Fig. 4(a). Details about the $[[4, 2, 2]]$ code and the distillation protocol as well as the data used to create Fig. 4(b) can be found in Appendix F.

Logical error rates in the Z , X , and Y bases (P_z , P_x , P_y), measured after a successful exit from the state preparation loop, are shown as black markers in Fig. 4(b). These are the fraction of trials in each basis for which the decoder inferred the incorrect logical state, normalized to the number of trials in which no uncorrectable error was detected. Decoding was attempted for trials with a single loss event, which can be corrected in the $[[4, 2, 2]]$ code. From these data, we extract a basis-averaged infidelity of $1 - F = 0.4^{(4)}_{(2)}\%$, using the definition $F = 1 - (P_x + P_y + P_z)/2$. We observe that the loop executes an average of 1.44 times before a correct value is measured, as shown in the histogram in Fig. 4(c).

TABLE I. Results for heralded fault-tolerant state preparation with conditional repetition, with result categories defined in text. Uncertainties represent 95% binomial confidence intervals.

Ancilla parity check	Loss conditioning	Success	Logical failure	Detected error
Success	All trials	$0.937^{(7)}_{(7)}$	$0.003^{(2)}_{(1)}$	$0.060^{(7)}_{(6)}$
	No loss	$0.887^{(9)}_{(9)}$	$0.002^{(2)}_{(1)}$	$0.11^{(1)}_{(1)}$
Failure	All trials	$0.65^{(2)}_{(2)}$	$0.055^{(1)}_{(1)}$	$0.30^{(2)}_{(2)}$
	No loss	$0.49^{(2)}_{(2)}$	$0.036^{(1)}_{(1)}$	$0.47^{(2)}_{(2)}$

The reset operation performed at the beginning of each state preparation attempt is implemented as an MCM cycle that contains an image of all qubits, so we can compare the logical failure rate of attempts where the ancilla did not have the correct parity (and the loop did not exit) to those where the correct parity was detected and the loop did exit. The red markers in Fig. 4(b) show the logical error rate for attempts where the incorrect parity was measured, with a basis-averaged infidelity of $11^{(3)}_{(2)}\%$.

The $[[4, 2, 2]]$ code can detect any single-qubit Pauli error, and can detect (but not correct) the presence of multiple losses. Trials where such errors are detected are not decoded, and are flagged as having a detected error. As an alternative way of decoding the data, we can condition the outputs on having no qubit loss, and only decode trials where no loss events were detected (labeling all loss events as a detected error). Table I shows the fraction of the total trials for which the output state was decoded to the correct state (“success”), for which it was decoded to the incorrect state (“logical failure”), and for which an error was detected that precluded the decoding of the state (“detected error”) for both analysis methods and results of the ancilla parity check.

Finally, we can compare the fidelities from the heralded distillation of logically encoded states to the fidelities measured from distillation using unencoded Bell states. We observe an unencoded fidelity of $97.7(3)\%$, compared to an encoded fidelity of $99.6^{(2)}_{(4)}\%$ showing superior performance for the encoded case.

VI. COHERENCE PRESERVATION DURING ATOM REPLENISHMENT

While we can refill lost atoms in the MZ with new atoms from the SZ, very long circuits will eventually deplete the population of the SZ, requiring replenishment to increase the logical lifetime beyond the lifetime of atoms in our system. Here, we refill the SZ with new atoms while maintaining coherence for atoms in the register. This capability is enabled by a combination of MCM,

as demonstrated above, as well as midcircuit conditional rearrangement and the ability to form a MOT in a spatially separated region from the tweezer array, as previously demonstrated in Ref. [13]. Additionally, midcircuit replenishment requires the ability to transport atoms from the MOT to tweezer regions and perform light-assisted collisions (LAC) that provide a sub-Poissonian distribution of 0 or 1 atom in each trap [26]. All operations must be performed while maintaining coherence of atoms in the register. The sequence used for replenishment is shown in Fig. 5(a), and described in detail in Appendix D.

We first load a core-shell-type MOT [27] with atoms from an atomic beam source. Atoms are transferred from the MOT into an optical lattice whose sites and focus translate to move the atoms near the tweezer arrays. Atoms are then loaded into the tweezers of the LZ, where LAC is performed. A block of imaging, cooling, and optical pumping using our standard MCM protocol allows the software service to identify populated sites within the LZ, as well as empty sites within the SZ. Atoms are then conditionally moved into the SZ sites, completing the replenishment sequence. The MOT may be loaded concurrently with other operations with negligible decoherence (see Appendix D). All other steps take approximately 300 ms to complete (the exact timing depends on the number of sites to be filled and the specific paths required for rearrangement).

In Fig. 5(b), we characterize the degree to which coherence is maintained in the register while replenishing atoms in the SZ. To do this, we perform a Ramsey sequence on atoms in the register with either zero hold time to extract background contrast, an idle period of equivalent length to a typical replenishment sequence, or a full replenishment sequence that refills the SZ. We analyze the remaining contrast in two ways: We can either average the Ramsey fringes of all sites together and then perform a global fit, or perform fits to single-site data and then average together the results. The former approach is more sensitive to spatially varying phase shifts than the latter. In this characterization, we focus on a single repetition of the replenishment sequence as this is how we anticipate the technique being used in a circuit, and to avoid complications interpreting the coherent addition or cancellation of errors that can occur with multiple sequential repetitions. With the global fit, we observe Ramsey contrasts of $99.8(4)\%$, $98.1(15)\%$, and $96.7(17)\%$ for the zero hold time, idle, and full sequence cases, respectively. With single-site fitting, we observe average Ramsey contrasts of $100.7(4)\%$, $99.2(5)\%$, $98.1(7)\%$ for the zero hold time, idle, and full sequence cases, respectively. The contrast values and uncertainties reported here are those returned from fits to the data, so can exceed unity due to statistical fluctuations (though above-unity contrast is of course unphysical). These results suggest that contrast loss during replenishment is largely due to spatially varying qubit phase shifts, due both to the

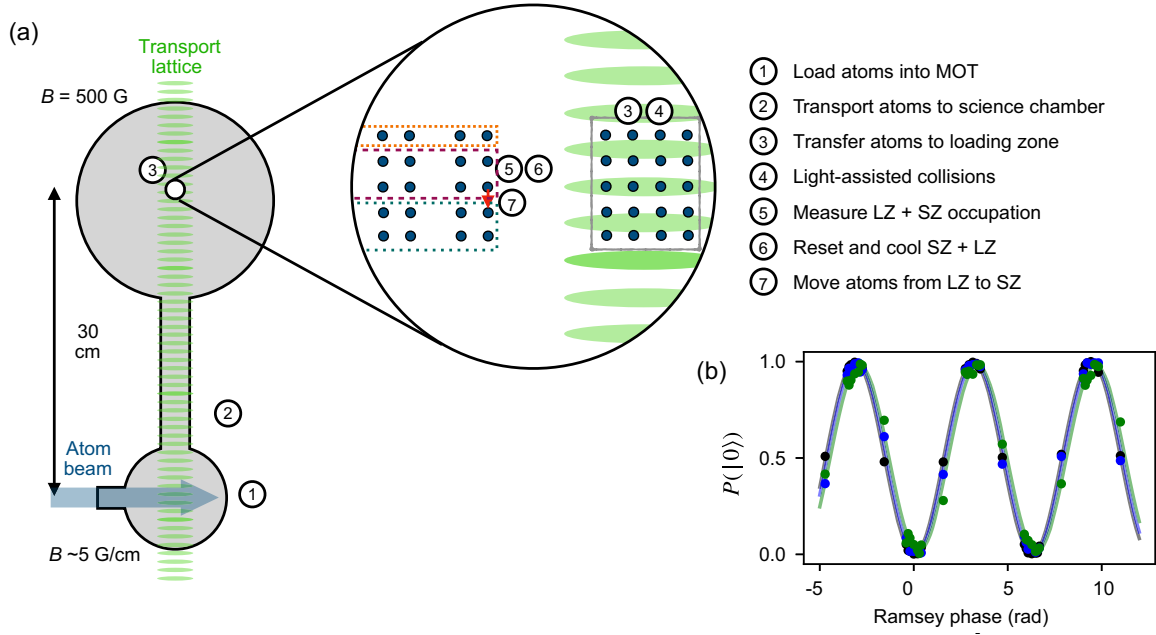


FIG. 5. Replenishing atoms in tweezer arrays. (a) Atoms are replenished into the storage zone from a magneto-optical trap (MOT), which is in turn filled from an atomic beam. The MOT is formed 30 cm away from the tweezer array, and then transported to the tweezer array using a moving optical lattice. After atoms are transferred from the lattice to LZ tweezers, light-assisted collisions (LACs) are induced to project tweezer populations to 0 or 1 atom. Occupation of the LZ and SZ are determined through an MCM block, after which atoms are moved from the LZ to the SZ to fill vacancies. All magnetic fields are static during all protocols reported here. (b) Coherence of atoms in the register is maintained during the atom replenishment process, as demonstrated by inserting the replenishment sequence inside a Ramsey sequence. Ramsey interference fringes, averaged over the full array, are shown for the case with the full replenishment sequence present (green), for an idle sequence of equal duration to the replenishment sequence (blue), and for a zero-hold-time sequence (black). The standard deviation of each point, calculated from a binomial distribution, is smaller than the markers.

lasers used in replenishment operations and to idle conditions in the trap (residual shifts from inhomogeneous trap depths and magnetic fields). These errors may potentially be mitigated with improved calibration or echo placement.

For our typical operating conditions, the fraction of SZ sites that are filled by the replenishment protocol exceeds 90%. Because the presence of atoms in the SZ is always measured before using those atoms to refill the MZ, the reloading fraction impacts only the frequency with which replenishment must be performed, but does not contribute directly to state preparation errors in a circuit.

VII. DISCUSSION AND CONCLUSIONS

In this work, we have demonstrated the ability to reinitialize and reuse ancilla atoms following a midcircuit measurement in a neutral atom quantum processor. By structuring error-correction cycles to periodically swap the role of data and ancilla qubits, all atoms in the array can be measured after a finite number of operations, allowing for extended error-correcting circuits to be run without accumulated errors due to atom heating and loss. We have also shown the ability to replenish atoms in the computational array from a thermal beam source midcircuit, in principle enabling the indefinite execution of quantum circuits on a platform where qubits have an inherently limited lifetime.

So far, we have used these capabilities to demonstrate a classical error-correction code—the repetition code. While the repetition code itself protects classical information (qubit populations, not coherences), the modified version we have demonstrated is also sensitive to potential phase errors during measurement, and so demonstrates the key capabilities needed for repeated execution of a true quantum error-correcting code (one that corrects both amplitude and phase errors) with atom reuse.

We have also shown how our MCM can be combined with real-time decision making to deterministically prepare a logical state heralded by the successful parity measurements of a block of ancilla qubits. In this case, the error rate on attempts with the correct ancilla parity is well below that of attempts with the incorrect ancilla parity, and below the error rate for unencoded distillation.

Finally, we have demonstrated the ability to replenish atoms in our tweezer arrays while maintaining coherence in existing atoms. This realizes a key step in solving a particular hurdle for neutral-atom-based quantum computation—the finite lifetime of atoms within the computational system—by in principle allowing the logical state of the system to persist beyond the lifetime of individual atoms. In this work, we have demonstrated this technique in isolation, rather than as a part of an

error-correcting circuit. For such applications, performance improvements beyond the numbers reported here would be desirable, and could likely be achieved using improved dynamical decoupling to better cancel coherent accumulation of errors, as well as by reducing the time required for rearrangement by improved control system latency, improved parallelism of moves, or better optimized move profiles.

DATA AVAILABILITY

The data supporting this study's findings are available within the article.

APPENDIX A: SYSTEM DETAILS

Atoms are loaded into the register and storage zone via the methods detailed in Ref. [13]. This protocol ensures a sufficient qubit count within the register for circuit execution and establishes an initial reservoir of atoms in the SZ to replenish atom loss observed during each MCM cycle. Typically, 128 register sites and 32 storage zone sites are filled. All zones, excluding the LZ, are arranged in eight pairs of matching columns, thereby creating highways for interzone atom transport. A static bias magnetic field of 500 G is aligned parallel to the array rows. A razor blade in an image plane of the register tweezers minimizes the 423-nm leakage light incident on any of the MZ, IZ, SZ, and LZ regions.

Atoms trapped in the 423-nm register tweezers have an average radial (axial) trap frequency of 56 kHz (12.5 kHz) during the initial loading, rearrangement, and 1Q gates. The atoms in IZ, MZ, and SZ have an average radial (axial) trap frequency of 70 kHz (15.5 kHz). The distance between neighboring columns of atoms is 3.3 μm [13], and the bottom row of MZ and the top row of register is separated by 33 μm , leaving 9 empty rows in between to facilitate row-wise atom movements and reduced 423-nm leakage light into IZ and MZ. The distance between pairs of columns (across the highways), which also corresponds to the distance between interaction pairs in the IZ, is 9.9 μm .

Arbitrary connectivity and selection of atoms for MCM is achieved by relocating selected atoms between different zones. Arbitrary 2Q gate connectivity is established by moving atoms between the register and the IZ, while MCM selection is achieved by moving selected atoms from the register to the MZ. Moves are performed in parallel between the register and other zones, up to specific constraints (see Ref. [10]). All movements within the LZ, MZ, and SZ in this work are executed by a single tweezer. These movements necessitate identifying a target and a source from different zones in real time, making these conditional movements. During any movement, atoms are transferred between static and rearrangement tweezers by ramping up the depth of the rearrangement tweezers

over 0.4 ms. We have not observed evidence of additional loss as a result of two-photon processes during handoff, which would manifest as a loss mechanism whose magnitude increases with trap powers. We have further confirmed that sums and differences of relevant trap photon energies do not coincide with known transitions from the ground state manifold, nor do two-photon transitions exceed the ionization threshold.

Our universal gate set is demonstrated in a static architecture in Ref. [16] and in a zoned architecture in Ref. [10]. We remark on three aspects of the current architecture. (1) 1Q gates are exclusively performed in the register. Atoms within a row are addressed in parallel, and different rows are addressed in series. Each qubit has 1Q gates applied at a single location within the register. See Ref. [10] and included references for a detailed description of the implementation of our 1Q gate scheme. (2) CZ gates are performed in parallel on up to eight pairs of arbitrary qubits, and (3) leakage into the $^3P_0$ and Rydberg states is converted to detectable atom loss after each 2Q gate with probabilities 99% and 80%, respectively, because the metastable states are not trapped in the 423-nm tweezers, and a fraction of Rydberg population decays to these states [10]. Because our imaging independently distinguishes the qubit state from atom loss [6], this allows for “delayed-erasure” decoding techniques to be applied [10–12].

APPENDIX B: MCM SEQUENCE

Determining both the state and presence of a qubit in the MZ requires two consecutive MCM images. The initial image identifies whether the atom is in the $|1\rangle$ state. To distinguish atoms in the $|0\rangle$ state from those that are lost, we collect a subsequent image after an optical pumping pulse that transfers all atoms to $|1\rangle$. We can also image the $|0\rangle$ state directly as demonstrated in Ref. [6], but performing two images of $|1\rangle$ allows us to cancel light shifts using spin-echo techniques, as described below.

The detailed beam geometry, level addressing, and sequence composition of the MCM block are shown in Fig. 6. Imaging is performed in a cavity-enhanced 784-nm 3D lattice by collecting the photons scattered from the cycling transition from state $|1\rangle$ to $^3P_1|F=3/2, m_F=3/2\rangle$ with a high numerical-aperture microscope objective [13]. The MCM image is 7 ms long, and the applied 556-nm light has a saturation parameter of approximately 1 and detuning of roughly 100 kHz. Because of the limited optical access, the imaging beam is inserted through one of the in-vacuum cavity mirrors with its k vector forming a 15-deg angle to the 500-G magnetic field, and illuminates the entire atom array. The beam is linearly polarized orthogonal to the magnetic field, having both σ^+ and σ^- components. Because of the Zeeman splitting of the $^3P_1$ sublevels, state-selective imaging of $|1\rangle$ is achieved using only the σ^+ component. Different frequency tones are generated by a fiber phase electro-optic modulator to reach

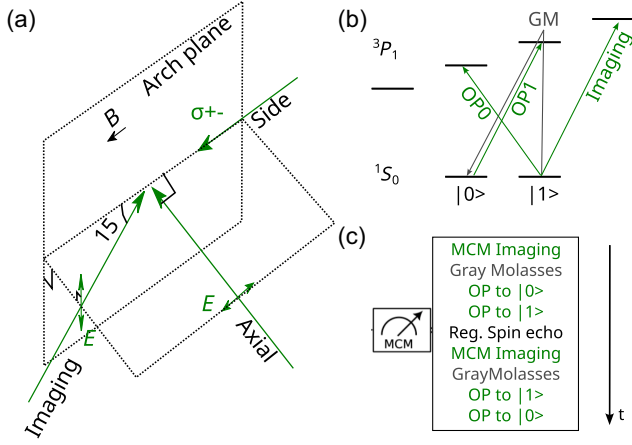


FIG. 6. (a) MCM beam geometry. “Arch” plane is the plane where the zoned architecture is formed. The “Imaging” beam addresses all zones. “Side” and “Axial” beams locally address IZ, SZ, and MZ. There are independent global “Side” and “Axial” beams for the regular (non-MCM) state preparation and measurement in the system (not shown). (b) MCM beam $^3P_1$ sublevel addressing scheme. Optical pumping to state $|0\rangle$ and $|1\rangle$ is provided by the “Side” beam. Gray molasses (GM) σ^+ and π are provided by the “Side” and “Axial,” respectively. (c) MCM block composition. The MCM sequence is symmetrized about a spin echo pulse in the register, though with the optical pumping order flipped.

each transition [6]. While the data presented in this work utilize only the σ^+ component directly, the presence of both components enables independent readout of both qubit states [6], which is a useful capability for calibrations and other protocols run on the same apparatus. Because the σ^- carrier is approximately 3 times farther detuned from any levels in $^3P_1$ compared to the σ^+ tone used to image $^3P_1|F=3/2, m_F=3/2\rangle$, and because we operate in a regime where the carrier is suppressed, scattering and loss from the σ^- component are subdominant.

Identification of the atomic state and presence is similar to cavity-based imaging presented in our previous works [10,13]. We identify the state and presence of atoms by comparing the number of photons collected in an image to a threshold. For MCM, care must be taken to avoid scattering of 423-nm leakage light into the MZ (see Fig. 7), which can degrade imaging performance. Measurement error information is shown in Table II. The MCM imaging incurs atom loss of 0.005(2) per image, higher than our non-MCM imaging due to optical pumping into $^3P_2$ state by 423-nm leakage light (see Appendix C). A 1388-nm $^3P_0$ repump laser is used for both regular and MCM imaging, and it improves the 784-nm cavity imaging loss by 0.0021(7) [0.0014(5)] for regular [MCM] imaging. For MCM images, we compromise state distinguishability slightly by reducing the number of imaging photons scattered in order to reduce atom loss. The spin flip probability from $|0\rangle$ to $|1\rangle$ is much lower than the spin flip probability from $|1\rangle$ to $|0\rangle$

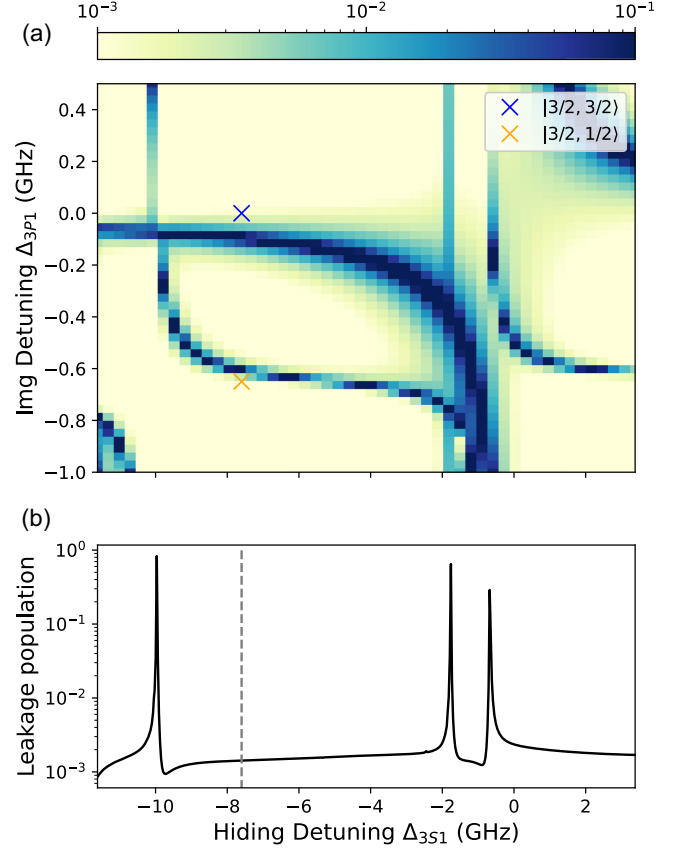


FIG. 7. Loss of register atoms during imaging. (a) Calculated per-image loss for our typical MCM imaging parameters (see text) for atoms in the register versus the register laser frequency (horizontal axis) and imaging laser frequency (vertical axis). Crosses represent our operating conditions for imaging (blue) and cooling (orange). Note that the scattering rate depicted here at the location of the orange cross is not relevant, as the polarization and intensity used during cooling are different from imaging. (b) A slice through the same data near our imaging detuning $\Delta_{3P1} = 0$. The dashed line represents our detuning Δ_{3S1} , indicating an expected loss near 10^{-3} for register atoms during an image. The resonances correspond to the different hyperfine levels of $8s^3S_1$.

because the $|0\rangle$ state is not excited to $^3P_1$. A background loss of about 0.0002 per image is incurred by our 30-s vacuum lifetime.

Following the MCM image, 4 ms of gray-molasses cooling is performed [10,19,28] by addressing the $^3P_1|3/2, 1/2\rangle$ state with two beams in a Raman configuration

TABLE II. Errors and loss during regular and MCM imaging.

	Regular imaging	MCM imaging
Atom loss per bright image	< 0.0006	0.005(2)
Distinguishability error	0.0003	0.003
$^3P_2$ loss/bright image	0.0004(6)	0.005(2)
$1 \rightarrow 0$		0.003(16)
$0 \rightarrow 1$		0.0006(8)

(σ^+ along the magnetic field and π through the objective) with a single-photon detuning of +1.5 MHz. The relative detuning of the two Raman beams is adjusted to be nearly equal to the qubit frequency splitting to give optimal cooling. Additionally, a 300- μ s optical pumping pulse prepares atoms in $|0\rangle$ ($|1\rangle$) by addressing the $^3P_1$ $|3/2, -1/2\rangle$ ($|3/2, 1/2\rangle$) states with a local σ -polarized beam (“Side” beam in Fig. 6). All beams used for cooling and optical pumping are localized to IZ, MZ, and SZ, though some leakage light into the register region is present. During the MCM cooling and optical pumping, atoms are simultaneously held in the optical cavity lattice and 423-nm tweezers, mitigating the effects of leakage light.

Given that our imaging protocol utilizes global 556-nm beams, qubits held within the register experience radian-scale phase shift as in Ref. [6]. To avoid the need to calibrate these phase shifts, and to cancel errors associated with any inhomogeneity or drift in these phase shifts, we implement 1Q BB1 composite echo pulses (described in [29]) on the qubits in the register. The echo pulse also cancels phase shifts due to the leakage of any local beams onto the register: After each MCM image, cooling and optical pumping are applied in order to ensure that any induced shifts cancel out. After the first image, optical pumping places atoms first into $|0\rangle$ and then into $|1\rangle$. After the second image, the order is reversed, preparing atoms in the SZ in the $|0\rangle$ state. The use of two optical pumping steps—one after each image—allows us to ensure that the same operations are applied before and after the echo pulse, helping us to cancel any phase shifts induced by these operations. Crucially, atoms in $|0\rangle$ are not promoted to the Rydberg state during subsequent 2Q gate operations, so atoms in the SZ do not interfere with gates in the nearby IZ.

After all atoms in both the MZ and the SZ are cooled and prepared in the $|0\rangle$ state, a software service identifies the locations of the remaining atoms in the SZ and lost atoms in the MZ. The service then streams real-time instructions to individually move the necessary atoms, correcting for detected loss events. The latency associated with this software service depends on the location and number of atoms to move, but for our typical circuits is consistently below 10 ms. We routinely achieve SZ-to-MZ conditional moves efficiencies exceeding 99.6%, measured through repeated cycles of populating one zone from the other using our standard MCM cycles.

APPENDIX C: LIGHT SHIFTS AND LOSS CHANNELS FROM REGISTER TWEEZER LIGHT

In order to shield atoms in the register from light used in the MCM sequence (either due to the global imaging beam or leakage from the local cooling beams), we operate the register tweezers near the $^3P_1$ - $8s^3S_1$ transition, which creates large shifts on the $^3P_1$ manifold. At our chosen

detuning, the shifts on the $^3P_1$ $|3/2, +3/2\rangle$ state (imaging) are approximately 2 times larger than the shifts on the $|3/2, \pm 1/2\rangle$ states (cooling and optical pumping). We increase the register power by 33% for the cooling step to balance the required shifts against the negative effects of light leakage. We also choose the optical power to avoid any accidental resonance product of the different frequency sidebands present after light is modulated [6].

We operate the register traps with a detuning of $\Delta_{3S1} \approx 2\pi \times -7.6$ GHz from the $^3P_1$ $|3/2, +3/2\rangle$ to $8s^3S_1$ $|3/2, +3/2\rangle$ transition with a Rabi frequency of $\Omega_{423} = 2\pi \times 2.4$ GHz during imaging. Although this provides an ac Stark shift of nearly 100 MHz on the imaged state, the combination of both the imaging light and register trap laser can cause the atoms to populate $8s^3S_1$ via a detuned two-photon process, which subsequently populates the long-lived ^3P states after spontaneous decay. This decay primarily populates the $^3P_2$ state, which is difficult to repump due to its large number of sublevels that are widely dispersed in energy by our large magnetic field. These atoms are eventually lost as our 459-nm traps do not trap the $^3P_2$ state. However, we are able to repump some small population from $^3P_0$, estimated to be less than 10% of the all leakage events caused by 423 nm.

To understand the scattering of the atoms during the hiding process, we construct a 16-level model including all hyperfine levels of $^1S_0$, $^3P_1$, and $8s^3S_1$, as well as additional leakage states to represent the long-lived ^3P states where the atoms are lost. We then numerically solve the optical Bloch equations for both the imaging and cooling processes assuming decay from the $^3P_1$ state to the ground state as well as decay from the $^3S_1$ state to both $^3P_1$ and the leakage state to calculate the population in the leakage state after these processes. A summary of the calculated populations during these different processes can be found in Table III. As shown in Fig. 7(a), the presence of the near-resonant 423-nm light shifts the resonances where the loss is enhanced by approximately -100 MHz (+50 MHz) for the $|3/2, 3/2\rangle$ ($|3/2, 1/2\rangle$), matching the measured Stark shift at this Rabi frequency. The crosses represent the operating points for the imaging and cooling, respectively, matching the

TABLE III. Summary of the sources and proportion of lost atoms during the MCM cycle. The calculated loss during the imaging and cooling processes accounts for the vast majority of observed loss in both the register and the MZ as measured in the main text.

	Register	MZ
Predicted imaging loss	0.003	0.005
Predicted cooling loss	0.006	0.00005
Predicted total loss	0.009	0.005
Measured total loss	0.0098(7)	0.005(2)

unshifted resonances for those states. In Fig. 7(b) we show a slice of the diagram at $\Delta_{3P1} = 0$, indicating that at our chosen detuning we will lose approximately 0.0015 of the register atoms during the imaging process. Given that each MCM cycle consists of two images, we expect this to contribute 0.003 to the total MCM loss for atoms in the register. The desired imaging tone is the first-order sideband generated by an electro-optic modulator (EOM). We have confirmed through our simulation that the light associated with the zeroth-order, second-order, and other first-order tones contributes scattering at least 2 orders of magnitude lower than the primary tone for our imaging conditions. However, the potential presence of additional rf spurs on the electronics used to drive the EOM could contribute additional loss.

Additional loss of atoms in the register is contributed by the cooling process. In particular, we find that loss in the register mostly depends on the overlap of the local σ^+ -polarized beam near the $|3/2, 1/2\rangle$ state that is used during cooling in the MZ and SZ. Simulations indicate that this polarization provides multiple loss channels through different intermediate states, including a transition to the $8s^3S_1$ state with less than 2 GHz of detuning. While this beam is nominally local to the measurement zone, we measure approximately 10% of the intensity incident on the register, leading to approximately a loss of 0.003 per cooling cycle (0.006 loss per MCM cycle). This loss channel in conjunction with the imaging loss described above accounts for the observed loss of register atoms described in the main text. Future work will improve the extinction ratio of the local beams to significantly reduce this effect.

Even small amounts of 423-nm light leaking into the MZ during MCM operations can lead to population of the higher-lying $^3S_1$ state that quickly decays to the longer-lived ^3P manifold. This is the main factor that leads to loss for atoms in the MZ during MCM cycles. This light leakage could be due to scattering in some of the many glass surfaces used in the optical path, as even a hard stop in an earlier image plane does not completely extinguish it.

We calculate that a extinction ratio of $1:10^{-5}$ (corresponding to an experimental bound on this leakage) for 423-nm light on the MZ can lead to an atom loss of 0.005 into the ^3P manifold after a $t_{\text{img}} = 7$ ms imaging pulse. This level of leakage light contributes a negligible Stark shift of less than 1% of the imaging Rabi frequency so does not appreciably modify the $^3P_1$ state population during imaging. However, this light off-resonantly drives atoms in the $^3P_1$ state to the leakage manifold via the $^3S_1$ state with a population proportional to $10^{-5}(\Omega_{423}/\Delta_{3S1})^2 \times \Gamma_{3S1}t_{\text{img}}$, leading to a loss rate of approximately 0.0025 per image (0.005 per MCM cycle). This can be reduced by increasing the detuning Δ or lowering the overall power of the tweezers to reduce $\Omega_{423,MZ}$, but at fixed contrast, this increases error rates on register atoms during MCM.

APPENDIX D: ATOM REPLENISHMENT SEQUENCE

The operation of our MOT and transport lattice are described in previous work [13], with the LZ referred to in this work corresponding to the reservoir of Ref. [13]. As before, to minimize 532-nm light incident on the data qubits stored in the register, the LZ is loaded with atoms from the transport lattice at a distance of 170 μm , and once loaded the whole array is moved near the other arrays using a galvo mirror in a conjugate plane. The LZ used in this work has 75 sites but no highways, such that the atom packing is denser (3.3- μm intersite).

We bound the decoherence contributed by operation of the MOT by measuring the decay in contrast in a Ramsey echo experiment with a MOT and without the MOT lasers on during the total hold time. Detailed timing information for our midcircuit reloading sequence is shown in Fig. 8. If a degradation of contrast occurs due to the formation of a MOT, we would expect it to have an exponential form versus time, as the most likely culprits would be associated with scattered light (the magnetic field environment does not change whether we form a MOT or not). From an exponential fit to contrast decay out to 2-s hold time, we observe a contrast decay rate of 0.029(3)/s and 0.035(4)/s without the MOT and with the MOT, respectively, corresponding to a differential contrast decay of 0.006(5)/s. This rate is negligible compared to other sources of error while running circuits, so the MOT can be loaded in parallel with other operations without meaningfully impacting performance. In contrast to Ref. [6], we use a core-shell MOT configuration [27], which drastically reduces 399-nm scattering from cooled atoms while maintaining a fast loading rate.

Given the low impact of the MOT during circuit execution (see below), we operate the MOT continuously in parallel with the circuits. The MOT is turned off only during the transport step while loading the atoms in the 532-nm conveyor belt lattice. The lattice takes 120 ms to transport the atoms 300 mm vertically from the MOT chamber to the science chamber. During the subsequent handoff, the lattice is ramped down over 10 ms while the loading tweezers are ramped up simultaneously. Finally, the galvo movement takes 10 ms approximately, leaving the array $\sim 6.6 \mu\text{m}$ from the SZ. The angular repeatability of our galvo (Thorlabs GVS211) is approximately 15 μrad , which translates to roughly 100 nm at the atom plane. We find that this level of repeatability is sufficient for reliable handoff to the movement tweezers. Because atoms are imaged and recooled after transfer to the computational arrays, small amounts of heating resulting from these fluctuations do not impact circuit performance, and we do not find calibration of the galvo position set point to be necessary in day-to-day operation.

Before rearranging atoms from the LZ to the SZ, we perform a parity projection step via LACs that creates

an array with single atoms and 50% loading fraction on average [26]. The LAC beams are local to the LZ, SZ, MZ, and IZ, and address the $^3P_1|3/2, \pm 1/2\rangle$ states. After single atoms are prepared, we use an MCM cycle to image and cool atoms in the LZ and SZ. Using these images, the software service calculates the rearrangement movements from the LZ to refill the SZ. The duration of these movements is nondeterministic as it depends on the number of sites to refill and its positions, but on average a full refilling of the SZ typically takes about 120 ms, including the software latency from the move compilation.

APPENDIX E: REPETITION CODE WITH ANCILLA REUSE

In a standard circuit-based repetition code memory experiment, the data qubits are not measured until the end of the circuit. This construction prevents the detection of qubit loss which is detected only when qubits are measured. Additionally, our atoms experience heating that scales with gate depth and movement count, so it is desirable to regularly image and reset the atoms to cool them.

To ensure that every atom is measured frequently, we make several changes from a standard repetition code.

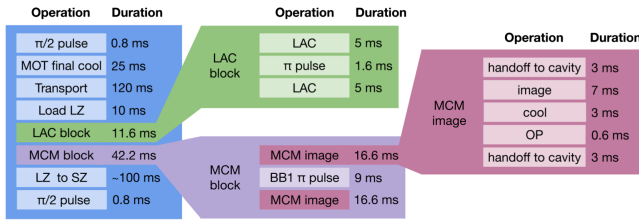


FIG. 8. Timing for midcircuit reloading sequence. See text in Appendix D for details. The $\pi/2$ pulses surrounding the reloading operations are included for characterization of decoherence caused by the reloading sequence. 1Q gates indicated are applied to atoms in the register.

First, we add an additional parity check, turning the repetition code into a 1D toric code, or ring code. Second, we compile SWAP gates into the syndrome extraction cycle, converting between syndrome qubits and data qubits every cycle [10–12]. After compilation to CZ gates, this SWAP operation adds no extra 2Q gates compared to a standard syndrome extraction circuit and requires only a possible Pauli correction that can be implemented in software. Furthermore, this change ensures that no atom remains active for more than four 2Q gates, in turn allowing us to detect lost atoms and cool those that remain. We implement SWAP operations so that the information stored in qubit i at the beginning of one cycle is transferred to qubit $i + 1 \bmod 2n$ at the end of the cycle. Thus, the quantum information gradually cycles around the ring; we refer to this circuit as a walking repetition code.

Following Refs. [20,25], we define “detectors” as sets of measurements with determined parity in the absence of error, conventionally defined as corresponding stabilizer measurements from one round to the next. With this convention, a given detector is sensitive only to errors in a localized region. For our walking repetition code, a parity check from measuring qubit j at the end of one cycle will correspond to a parity check measured on qubit $j + 1 \bmod 2n$ at the end of the subsequent cycle. The set of two such measurements (in the bulk) defines a detector, with expected even parity. In the first cycle and for the final measurement, we define similar detectors with known parity by comparing to known initial or final states. Defining detectors in this manner creates localized regions of error sensitivity, as shown in Fig. 9. These regions can be computed by propagating Pauli errors that anticommute with stabilizer measurements, following the Pauli web rules in Ref. [30]. To compute detector values in the presence of atom loss, we populate bits corresponding to lost atoms with random values.

Since the repetition code protects only classical information, the detection regions are not sensitive to all

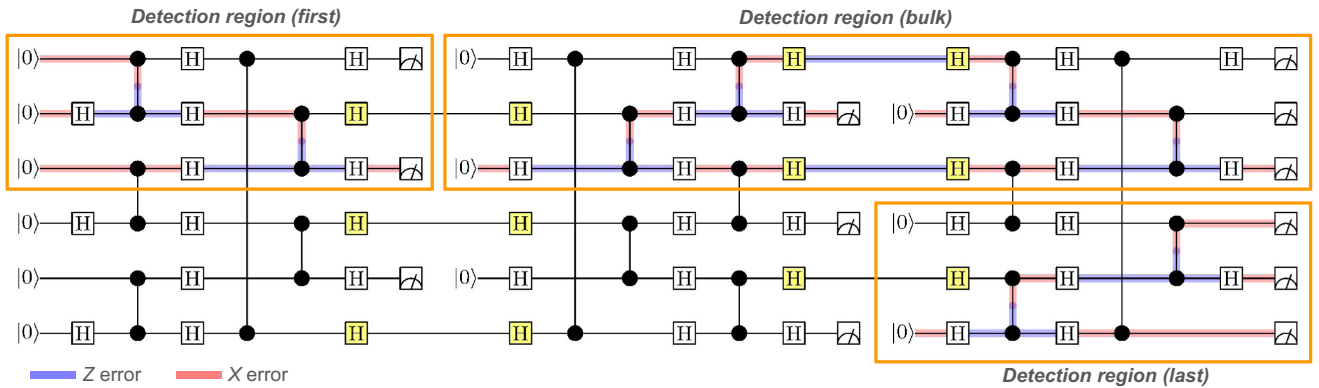


FIG. 9. Example detection regions for a distance-3 phase-sensitive repetition code, indicating the locations in the circuit where an error will lead to a specific detector event. We show sensitivity for both Z-type (blue) and X-type (red) Pauli errors. Any single Pauli error of appropriate type changes the parity of the corresponding detector. The phase-insensitive version of the circuit excludes the Hadamard gates (highlighted) on idle qubits surrounding the MCMs.

TABLE IV. Decoding failure data for repetition code experiment runs.

Distance	Logical failures	Total shots
Phase sensitive		
3	30	11700
5	11	11250
7	6	11250
Phase insensitive		
3	25	17820
5	3	17640
7	4	17460
9	3	16920

quantum channels. Since we anticipate Z-biased noise on idle qubits during MCM, we construct two different versions of the repetition code circuits, a phase-insensitive version that does not detect phase errors during MCM and a phase-sensitive version that detects phase errors during MCM. As expected, we see increased detection frequency for the phase-sensitive version as discussed in the main text.

In order to decode these results, we used the sparse-blossom variant of the minimum-weight perfect matching algorithm, implemented by `PyMatching` [24]. The initial matching graph was generated using `Stim` based on a predetermined error model [31]. The noise parameters used in this model were initially based on fidelity estimates and gate characterizations, and were then fine-tuned over earlier runs of repetition code experiments. For the data shown in this paper, we used the final set of fine-tuned parameters.

To account for atom loss, the matching graph was edited on a per-shot basis, using techniques similar to those described in Ref. [32]. For each detected atom loss, we modify the edge weights in the matching graph in two ways. First, we update timelike edges (corresponding to readout errors) to have zero weight when connected to detectors affected by atom loss. Second, we modify edge weights corresponding to Pauli probabilities on partner qubits sharing CZ gates with a potentially lost atom. We call these errors “loss-correlated gate errors.” A comparison between various decoders will be discussed in future work.

In Table IV, we provide the data used to generate Fig. 3(c).

APPENDIX F: ENCODED BELL PAIR DISTILLATION

The $[[2, 1, 2]]$ distillation protocol [33] is illustrated in Fig. 10. To detect bit-flip errors during preparation of a Bell pair, a second Bell pair is also prepared. A transversal CNOT gate is applied between the two Bell pairs, propagating bit-flip errors from the control block to the target block. Measuring the target block in the incorrect state (odd parity) indicates an error during the preparation of one of the Bell pairs. This information is insufficient to determine

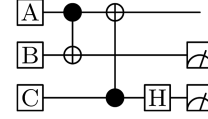


FIG. 10. Schematic of the $[[2, 1, 2]]$ distillation protocol. Blocks A, B, and C represent Bell pairs, either encoded or unencoded. The CNOT gates are applied in a transversal manner between the blocks. The procedure is complete if both blocks B and C are measured in even parity states.

where the error occurred, so if the measurement indicates an error, both blocks are discarded. This procedure suppresses physical errors at first order. An analogous procedure can be applied to detect phase errors.

When operating on physical Bell pairs, this procedure is not tolerant to gate faults. To make the procedure fault tolerant requires encoding the procedure into logical blocks. In this work, we prepare logical Bell pairs using $[[4, 2, 2]]$ codes [34–36]. This code is a stabilizer code [37] that can detect (but not correct) any single-qubit Pauli error and can correct for any single-qubit loss event. The $[[4, 2, 2]]$ code stores two logical qubits in four physical qubits and admits a transversal CNOT operation that acts as logical CNOT gates between corresponding logical qubits across two $[[4, 2, 2]]$ code blocks. When analyzing final logical states, we call a detected stabilizer violation a “detected error,” while a “logical failure” corresponds to the correct measured stabilizer, but the incorrect logical value.

To implement an encoded distillation scheme, we prepare three $[[4, 2, 2]]$ code blocks in logical Bell states. The preparation of physical states $|\Phi^+\rangle \otimes |\Phi^+\rangle$ fault-tolerantly prepares an undistilled logical Bell pair $|\tilde{\Phi}^+\rangle$ in the $[[4, 2, 2]]$ code, as demonstrated previously in Ref. [10]. We treat one block as the “data block” while the other two blocks are used to detect bit-flip and phase errors. The full protocol prepares logical Bell states in a manner that suppresses any single Pauli error. After the ancillary blocks are measured using MCM, the same software service that deals with conditional movements between zones decodes the ancilla measurement. Following this check, we rotate the data qubits in order to measure the logical qubits in the XX , YY , or ZZ basis. The loop exit is triggered by a successful check, after which the data qubits are measured in the register. When decoding the ancilla block, a “success” is defined as the correct parity measurement with all atoms present. If the state preparation is not successful, all physical qubits are reset via another MCM block and the procedure starts over. Note that the final measurement of this procedure reintroduces errors that are detectable but not suppressed by the distillation procedure.

In addition to the procedure described above, we performed a variant in which antiferromagnetic states were created. This corresponds to the preparation of entangled states $|\Psi^-\rangle$. Two copies of this physical state produce an encoded state $|\tilde{\Psi}^-\rangle$. These states are insensitive to global

TABLE V. Measurement results for unencoded and encoded Bell pair distillation. Here, a failure corresponds to an incorrect parity measurement for the target Bell state.

	XX basis	YY basis	ZZ basis
	Unencoded		
Successes	2443	2127	4828
Failures	7	61	79
	Encoded		
Successes	1440	1366	1452
Failures	3	3	4

coherent phase noise, even when physical qubits are rotated using Hadamard gates. We do not observe any significant difference when performing these experiments, so we do not distinguish among them in the analysis shown here.

Here we include the raw experimental data used to calculate the fidelities for encoded and unencoded distillation experiments reported in the main text in Table V.

-
- [1] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, *Phys. Rev. A* **52**, R2493 (1995).
- [2] E. Knill and R. Laflamme, *Theory of quantum error-correcting codes*, *Phys. Rev. A* **55**, 900 (1997).
- [3] T. M. Graham, L. Phuttitarn, R. Chinnarasu, Y. Song, C. Poole, K. Jooya, J. Scott, A. Scott, P. Eichler, and M. Saffman, *Midcircuit measurements on a single-species neutral alkali atom quantum processor*, *Phys. Rev. X* **13**, 041051 (2023).
- [4] E. Deist, Y.-H. Lu, J. Ho, M. K. Pasha, J. Zeiher, Z. Yan, and D. M. Stamper-Kurn, *Mid-circuit cavity measurement in a neutral atom array*, *Phys. Rev. Lett.* **129**, 203602 (2022).
- [5] K. Singh, C. Bradley, S. Anand, V. Ramesh, R. White, and H. Bernien, *Mid-circuit correction of correlated phase errors using an array of spectator qubits*, *Science* **380**, 1265 (2023).
- [6] M. A. Norcia *et al.*, *Midcircuit qubit measurement and rearrangement in a ^{171}Yb atomic array*, *Phys. Rev. X* **13**, 041034 (2023).
- [7] J. W. Lis, A. Senoo, W. F. McGrew, F. Rönchen, A. Jenkins, and A. M. Kaufman, *Midcircuit operations using the omg architecture in neutral atom arrays*, *Phys. Rev. X* **13**, 041035 (2023).
- [8] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz *et al.*, *Logical quantum processor based on reconfigurable atom arrays*, *Nature (London)* **626**, 58 (2024).
- [9] R. Finkelstein, R. B.-S. Tsai, X. Sun, P. Scholl, S. Direkci, T. Gefen, J. Choi, A. L. Shaw, and M. Endres, *Universal quantum operations and ancilla-based read-out for tweezer clocks*, *Nature (London)* **634**, 321 (2024).
- [10] B. W. Reichardt, A. Paetznick, D. Aasen, I. Basov, J. M. Bello-Rivas *et al.*, *Logical computation demonstrated with a neutral atom quantum processor*, *arXiv:2411.11822*.
- [11] M. N. H. Chow, V. Buchemmavari, S. Omanakuttan, B. J. Little, S. Pandey, I. H. Deutsch, and Y.-Y. Jau, *Circuit-based leakage-to-erasure conversion in a neutral-atom quantum processor*, *PRX Quantum* **5**, 040343 (2024).
- [12] H. Perrin, S. Jandura, and G. Pupillo, *Quantum error correction resilient against atom loss*, *Quantum* **9**, 1884 (2025).
- [13] M. A. Norcia, H. Kim, W. B. Cairncross, M. Stone, A. Ryou, M. Jaffe, M. O. Brown *et al.*, *Iterative assembly of ^{171}Yb atom arrays with cavity-enhanced optical lattices*, *PRX Quantum* **5**, 030316 (2024).
- [14] W. Huie, L. Li, N. Chen, X. Hu, Z. Jia, Won Kyu Calvin Sun, and J. P. Covey, *Repetitive readout and real-time control of nuclear spin qubits in ^{171}Yb atoms*, *PRX Quantum* **4**, 030337 (2023).
- [15] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, *Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays*, *Nat. Commun.* **13**, 4657 (2022).
- [16] J. A. Muniz, M. Stone, D. T. Stack, M. Jaffe, J. M. Kindem *et al.*, *High-fidelity universal gates in the ^{171}Yb ground-state nuclear-spin qubit*, *PRX Quantum* **6**, 020334 (2025).
- [17] K. Barnes, P. Battaglini, B. J. Bloom, K. Cassella, R. Cox, N. Crisosto, J. P. King, S. S. Kondov, K. Kotru, S. C. Larsen *et al.*, *Assembly and coherent control of a register of nuclear spin qubits*, *Nat. Commun.* **13**, 2779 (2022).
- [18] T. O. Höhn, E. Staub, G. Brochier, N. Darkwah Oppong, and M. Aidelburger, *State-dependent potentials for the $^1\text{S}_0$ and $^3\text{P}_0$ clock states of neutral ytterbium atoms*, *Phys. Rev. A* **108**, 053325 (2023).
- [19] A. Jenkins, J. W. Lis, A. Senoo, W. F. McGrew, and A. M. Kaufman, *Ytterbium nuclear-spin qubits in an optical tweezer array*, *Phys. Rev. X* **12**, 021027 (2022).
- [20] Z. Chen, K. J. Satzinger *et al.*, *Exponential suppression of bit or phase errors with cyclic error correction*, *Nature (London)* **595**, 383 (2021).
- [21] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, H. Levine, G. Semeghini, M. Greiner, V. Vuletić, and M. D. Lukin, *High-fidelity parallel entangling gates on a neutral-atom quantum computer*, *Nature (London)* **622**, 268 (2023).
- [22] D. Cruz, R. Fournier, F. Gremion, A. Jeannerot, K. Komagata, T. Tosić, J. Thiesbrummel, C. L. Chan, N. Macris, M.-A. Dupertuis *et al.*, *Efficient quantum algorithms for GHZ and W states, and implementation on the IBM quantum computer*, *Adv. Quantum Technol.* **2**, 1900015 (2019).
- [23] S. A. Moses, C. H. Baldwin, M. S. Allman *et al.*, *A race-track trapped-ion quantum processor*, *Phys. Rev. X* **13**, 041052 (2023).
- [24] O. Higgott and C. Gidney, *Sparse blossom: Correcting a million errors per core second with minimum-weight matching*, *Quantum* **9**, 1600 (2025).
- [25] A. G. Fowler, D. Sank, J. Kelly, R. Barends, and J. M. Martinis, *Scalable extraction of error models from the output of error detection circuits*, *arXiv:1405.1454*.
- [26] N. Schlosser, G. Reymond, I. Protchenko, and P. Grangier, *Sub-Poissonian loading of single atoms in a microscopic dipole trap*, *Nature (London)* **411**, 1024 (2001).
- [27] J. Lee, J. H. Lee, J. Noh, and J. Mun, *Core-shell magneto-optical trap for alkaline-earth-metal-like atoms*, *Phys. Rev. A* **91**, 053405 (2015).

- [28] L. Li, X. Hu, Z. Jia, W. Huie, W. K. C. Sun, Aakash, Y. Dong, N. Hiri-O-Tuppa, and J. P. Covey, *Parallelized telecom quantum networking with an ytterbium-171 atom array*, Nat. Phys. (2025), [10.1038/s41567-025-03022-4](https://doi.org/10.1038/s41567-025-03022-4).
- [29] S. Wimperis, *Broadband, narrowband, and passband composite pulses for use in advanced NMR experiments*, J. Magn. Reson., Ser. A **109**, 221 (1994).
- [30] H. Bombin, D. Litinski, N. Nickerson, F. Pastawski, and S. Roberts, *Unifying flavors of fault tolerance with the ZX calculus*, Quantum **8**, 1379 (2024).
- [31] C. Gidney, *stim: A fast stabilizer circuit simulator*, Quantum **5**, 497 (2021).
- [32] G. Baranes, M. Cain, J. Ataiades, D. Bluvstein, J. Sinclair, V. Vuletic, H. Zhou, and M. D. Lukin, *Leveraging atom loss errors in fault tolerant quantum algorithms*, [arXiv:2502.20558](https://arxiv.org/abs/2502.20558).
- [33] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, *Quantum privacy amplification and the security of quantum cryptography over noisy channels*, Phys. Rev. Lett. **77**, 2818 (1996).
- [34] L. Vaidman, L. Goldenberg, and S. Wiesner, *Error prevention scheme with four particles*, Phys. Rev. A **54**, R1745 (1996).
- [35] M. Grassl, T. Beth, and T. Pellizzari, *Codes for the quantum erasure channel*, Phys. Rev. A **56**, 33 (1997).
- [36] N. M. Linke, M. Gutierrez, K. A. Landsman, C. Figgatt, S. Debnath, K. R. Brown, and C. Monroe, *Fault-tolerant quantum error detection*, Sci. Adv. **3**, e1701074 (2017).
- [37] D. Gottesman, *Theory of fault-tolerant quantum computation*, Phys. Rev. A **57**, 127 (1998).