Fault-Tolerant Operation of a Quantum Error-Correction Code

Laird Egan[1,†], Dripto M. Debroy[2], Crystal Noel[1], Andrew Risinger[1], Daiwei Zhu[1], Debopriyo Biswas[1], Michael Newman[3,*], Muyuan Li[5], Kenneth R. Brown[2,3,4,5], Marko Cetina[1,2], and Christopher Monroe[1]

[1]Joint Quantum Institute, Center for Quantum Information and Computer Science, and Departments of Physics and Electrical and Computer Engineering, University of Maryland, College Park, MD 20742
[2]Department of Physics, Duke University, Durham, NC 27708
[3]Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708
[4]Department of Chemistry, Duke University, Durham, NC 27708
[5]Schools of Chemistry and Biochemistry and Computational Science and Engineering, Georgia Institute of Technology, Atlanta, GA, 30332
[*]Present Address: Google Research, Venice, CA 90291
[†]Email: lne@umd.edu

ABSTRACT

Quantum error correction protects fragile quantum information by encoding it in a larger quantum system whose extra degrees of freedom enable the detection and correction of errors. An encoded logical qubit thus carries increased complexity compared to a bare physical qubit. Fault-tolerant protocols contain the spread of errors and are essential for realizing error suppression with an error-corrected logical qubit. Here we experimentally demonstrate fault-tolerant preparation, rotation, error syndrome extraction, and measurement on a logical qubit encoded in the 9-qubit Bacon-Shor code. For the logical qubit, we measure an average fault-tolerant preparation and measurement error of 0.6% and a transversal Clifford gate with an error of 0.3% after error correction. The result is an encoded logical qubit whose logical fidelity exceeds the fidelity of the entangling operations used to create it. We compare these operations with non-fault-tolerant protocols capable of generating arbitrary logical states, and observe the expected increase in error. We directly measure the four Bacon-Shor stabilizer generators and are able to detect single qubit Pauli errors. These results show that fault-tolerant quantum systems are currently capable of logical primitives with error rates lower than their constituent parts. With the future addition of intermediate measurements, the full power of scalable quantum error-correction can be achieved.

Quantum computers promise to solve models of important physical processes, optimize complex cost functions, and challenge cryptography in ways that are intractable using current computers[1–5]. However, realistic quantum component failure rates are typically too high to achieve these goals[6,7]. These applications will therefore likely require quantum error correction schemes to significantly suppress errors[8,9].

Quantum error correcting codes combine multiple physical qubits into *logical* qubits that robustly store information within an entangled state[10–12]. However, these codes are not enough on their own. Fault-tolerant (FT) logical operations, which limit the ways in which errors can spread throughout the system, must also be used. Without them, the logical error rate will remain proportional to the physical error rate. FT state preparation, detection, and operations have been demonstrated using quantum error detecting codes with four data qubits[13–17]. These codes can identify when errors have occurred, but do not extract enough information to correct them. There have also been quantum demonstrations of classical repetition codes to correct quantum errors restricted along one axis [18–23]. In other work, qubits have been encoded into quantum error correcting codes that can correct all single qubit errors, but the encoding procedure was not fault tolerant[24] and the system was not large enough to measure all the error syndromes[25,26]. Parallel work on bosonic codes has demonstrated encoded operations[27,28], fault-tolerant detection, one-axis[29], and two-axis[30] error correction on encoded qubits, but has yet to demonstrate fault-tolerant state encoding[31]. For both qubit codes and bosonic codes, fault-tolerant state preparation of a code capable of correcting all single-qubit errors has not been achieved.

Here, we present the first implementation of FT logical state preparation. Remarkably, the encoding outperforms the limiting physical operation, which, in our system, is the two-qubit entangling operation. We integrate the encoded logical qubit with FT stabilizer measurement, FT logical gates, and FT measurement to achieve a code capable of correcting errors along all axes. We achieve these results through the first experimentally implemented subsystem quantum error correction code. In addition, we demonstrate non-FT direct preparation and continuous logical gates for the creation of arbitrary logical states and compare these to fault-tolerant protocols. In the process, we generate high-quality encoded magic states, which are a critical resource for many universally programmable FT quantum computing architectures[32]. Achieving all these subroutines in a single system establishes a key landmark for fault-tolerant quantum computing.

The quantum computer used in this work consists of laser-

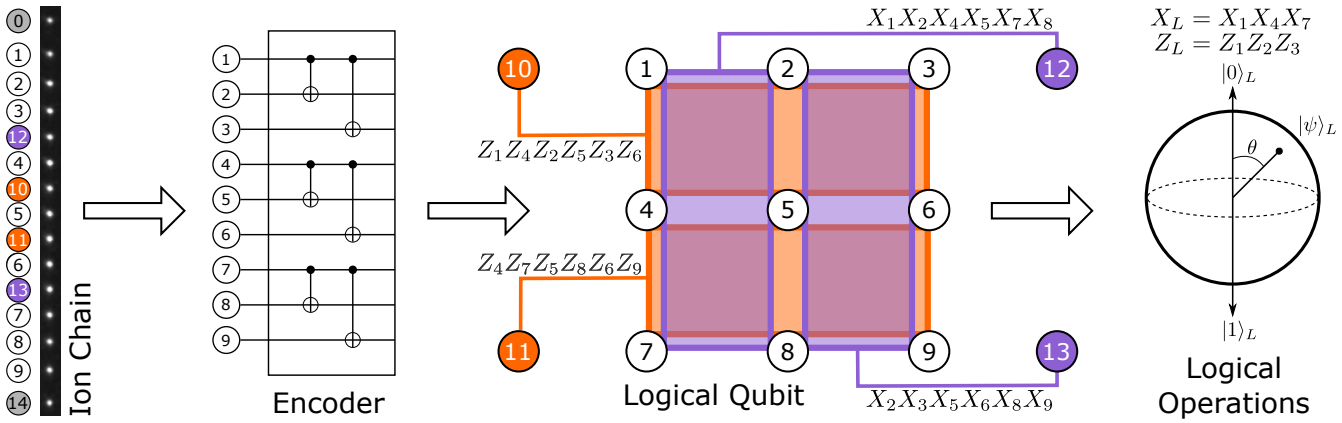

Figure 1. **The Bacon-Shor subsystem code implemented on a 15 ion chain.** Bacon-Shor is a [[9,1,3]] subsystem code that encodes 9 data qubits into 1 logical qubit. Four weight-6 stabilizers are mapped to ancillary qubits 10, 11, 12, and 13, for measuring errors in the X and Z basis. We demonstrate encoding of the logical qubit, with subsequent stabilizer measurements and logical gate operations.

cooled $^{171}Yb^+$ ions trapped above a microfabricated chip[33] in a room-temperature vacuum chamber. Each physical qubit is encoded in the $^2S_{1/2}$ electronic ground state hyperfine 'clock' states of a single $^{171}Yb^+$ ion, $|0\rangle \equiv |F=0;m_F=0\rangle$, $|1\rangle \equiv |F=1;m_F=0\rangle$, with a qubit frequency splitting of $\omega_0 = 2\pi \times 12.642820424(4)$ GHz. The qubits have a measured T_2 decoherence time in excess of 2.75 s (limited by the stability of external magnetic fields) and average single-shot detection fidelity of $> 99.5\%$. Quantum gates are driven by individually optically addressing up to 32 equispaced ions in a single chain via a multi-channel acousto-optic modulator (AOM)[34]. We implement high-fidelity native single-qubit and two-qubit gates with fidelities of 99.98% and 98.5-99.3%, respectively. All-to-all two-qubit gate connectivity is achieved through coupling of ions via a shared motional bus[35].

As shown in Fig. 1, we implement a [[9,1,3]] Bacon-Shor code [36,37], which is well-suited to near-term ion-trap quantum computing architectures[38]. This code uses 9 physical qubits to encode 1 logical qubit to distance 3, meaning at least 3 single-qubit operations are needed to change the logical state. Stabilizers can be measured using 4 ancilla qubits[39]. The Bacon-Shor code is a subsystem code with 4 additional degrees of freedom, known as gauge qubits, which can be used for designing stabilizer measurement circuits with minimal resources. The Bacon-Shor code is a generalization of Shor's code[11], which is a concatenation of bit-flip and phase-flip repetition codes. Its logical states are products of GHZ states:

$$\begin{aligned} |0/1\rangle_L \otimes |\overline{X}\rangle_G &= \frac{1}{2\sqrt{2}}(|{+}{+}{+}\rangle \pm |{-}{-}{-}\rangle)^{\otimes 3}, \\ |+/-\rangle_L \otimes |\overline{Z}\rangle_G &= \frac{1}{2\sqrt{2}}(|000\rangle \pm |111\rangle)^{\otimes 3}, \end{aligned} \tag{1}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|\overline{X}/\overline{Z}\rangle_G$ refer to different states of the gauge qubits; see Supplementary Information for details.

Bacon-Shor codes support a wide range of FT operations, including state preparation, stabilizer measurement, state measurement, and gates. Fault-tolerance, as a design principle, ensures faults on single operations do not propagate to uncorrectable multi-qubit failures in the circuit. When encoding a logical qubit, entanglement must be generated between the data qubits. In most quantum error correcting codes, this is done by repeatedly measuring stabilizers in order to project into the code subspace, which requires additional ancilla qubits and entangling gate operations. Unlike many other codes (e.g., topological), not all Bacon-Shor logical states require global entanglement, as seen in Eq. 1. These states can be structured so that any single circuit fault during preparation can result in, at worst, a global phase after correction. Thus, Bacon-Shor codes are among the few code families that can be fault-tolerantly prepared with unitary operations. Stabilizer measurement also requires interacting ancillae with multiple data qubits, which could cause damaging correlated errors; however, by carefully ordering the gates, the correlated errors affect only the gauge qubits while leaving the logical qubit untouched[39,40]. FT state measurement can be performed by simply measuring the data qubits followed by classical error correction on the output.

Fault tolerance in logical gates is often achieved through *tranversality*, where there are no interactions between data qubits in the same code block. When including permutations, Bacon-Shor has transversal constructions for $\{CNOT_L, H_L, Y_L(\pi/2), X_L\}$, along with X_L and Z_L basis preparation[41–43]. Here, $Y(\theta)$ indicates exponentiation of the Pauli-$\bar{Y}$ matrix, $e^{-i\theta\bar{Y}/2}$. In this work we also implement a non-FT logical $Y_L(\theta) = Y_1X_2X_3Z_4Z_7(\theta)$ gate and a non-FT $|\psi\rangle_L = \cos(\theta/2)|0\rangle_L - i\sin(\theta/2)|1\rangle_L$ preparation circuit for the construction of arbitrary logical states. The latter demonstrates generation of high-fidelity magic states, which can be used to fault-tolerantly implement non-Clifford gates[32].

Encoding the Logical Qubit

We embed the 9 data qubits and 4 ancilla qubits of the Bacon-Shor-13 code in a single chain of 15 ions (Fig. 1), with the

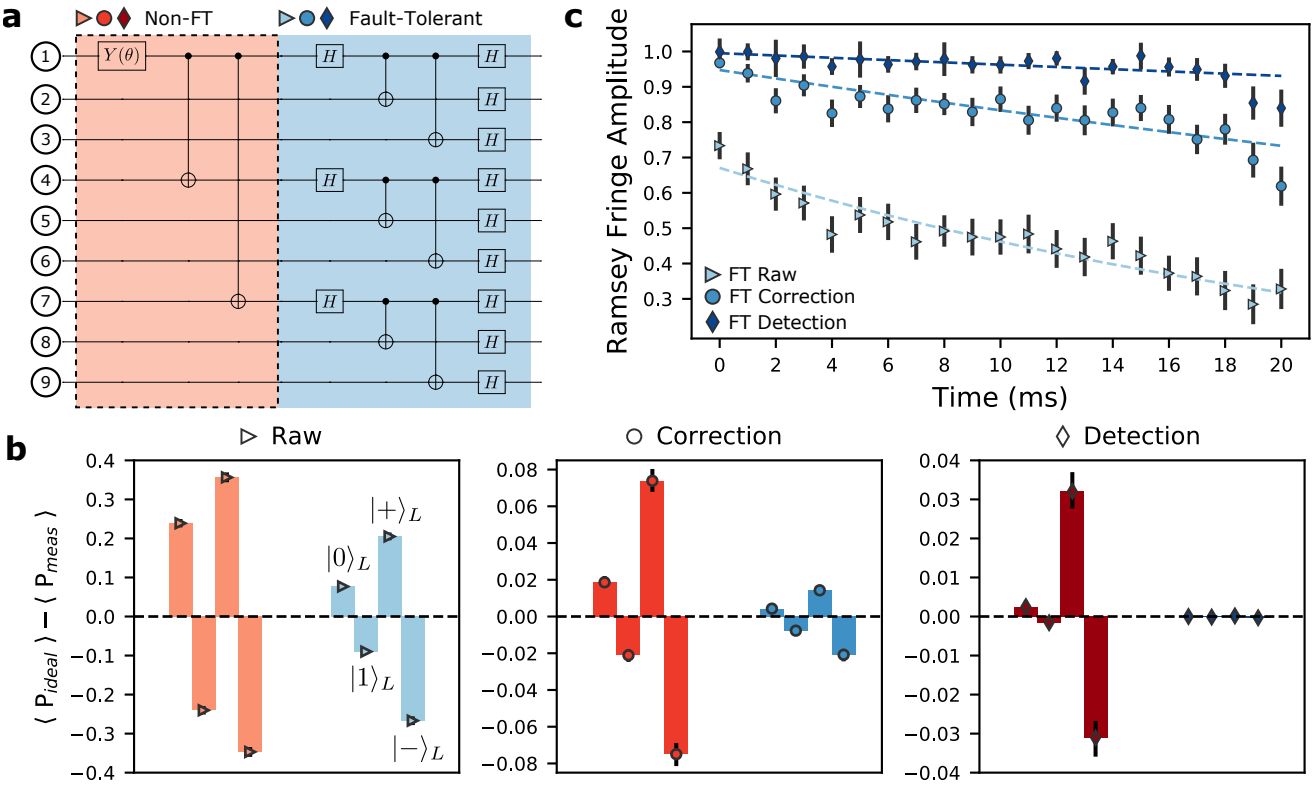

Figure 2. **Fault-tolerant logical qubit state preparation. a,** Encoding circuit for creating logical qubit basis states. The right subcircuit (blue) is used for FT preparation of Z/X logical basis states. The left subcircuit (red, dashed) can be optionally prepended for non-FT preparation of arbitrary logical states in the xz-plane. **b,** Bit flip errors for the key basis states of the encoded logical qubit. $|0/1\rangle_L$ states are directly measured in the Z-basis while the $|+/-\rangle_L$ states are measured in the X basis by applying a transversal $Y_L(-\pi/2)$ after encoding. The measured expectation value of the parity ($P_{meas} = \langle Z/X \rangle_L$) is compared against the ideal parity of the logical state ($P_{ideal} = \pm 1$). Error bars indicate the 95% binomial proportion confidence interval. **c,** Phase flip errors as a function of time measured for the $|+\rangle_L$ state. After each wait time, a varying $Z(\phi)$ gate is applied to every data qubit, followed by $Y_L(-\pi/2)$. A fit of $\langle X \rangle_L$ depending on ϕ to a Ramsey fringe yields the Ramsey amplitude. Error bars are the 95% confidence intervals from maximum likelihood estimation fits.

two end ions left idle to obtain uniform spacing of the central 13 ions. The mapping of the code onto the chain is chosen to minimize two-qubit gate crosstalk (details in Supplementary Information).

The qubits are initialized to the $|0\rangle$ state by optical pumping. We perform quantum gate operations with 355 nm laser beams that address the target qubits via an optical stimulated Raman process. A "global" beam illuminates all the ions and is perpendicular to an array of 32 tightly-focused "individual" beams, which are independently controlled via a multichannel AOM. By sending appropriate signals to the AOM, we can execute universal single and two-qubit gates on the ion chain with all-to-all connectivity[34,35]. At the end of a circuit, we perform global state readout by simultaneously collecting state-dependent fluorescence from each ion using high-resolution optics and 32 individual photo-multiplier tubes.

The encoding circuit used to create logical states is shown in Fig. 2(a). The right sub-circuit (blue) is FT because there are no entangling operations between independent GHZ states that would allow errors to propagate; however it is limited to preparation of only X and Z basis states. One may prepend an optional sub-circuit (red, dashed) that enables the encoding of arbitrary $|\psi\rangle_L$ states in the xz-plane, determined by the initial single qubit rotation $Y(\theta)$. This circuit can produce global entanglement, and allows the possibility of early errors spreading between the separate GHZ states. As a consequence, this circuit loses the FT properties of the X and Z basis preparation circuits. To directly investigate the properties of fault-tolerance, we compare the encoding performance of the right sub-circuit to the full circuit with $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$.

Logical qubit states are measured by looking at the expected value of the total parity of all the data qubits in the Z-basis, $\langle Z \rangle_L = \langle Z_1 Z_2 ... Z_8 Z_9 \rangle$. From Eq.1, the $|0\rangle_L$ state has even parity ($\langle Z \rangle_L = 1$) while $|1\rangle_L$ has odd parity ($\langle Z \rangle_L = -1$). Similarly, the $|+/-\rangle_L$ states have even/odd parity after a $Y_L(-\pi/2)$ operation following the encoding circuit, which maps $\langle X \rangle_L \rightarrow \langle Z \rangle_L$. The measured parity compared to the ideal parity of each logical Z, X basis state is presented in Fig. 2(b). The three panels correspond to different postprocessing techniques applied to the raw measurement output.

The raw panel corresponds to simply calculating the parity of all the data bits ($\langle Z\rangle_L$). As such, a single bit-flip error on the data can cause this decoding method to fail. The correction panel corresponds to applying error-correction by flipping the sign of the logical qubit parity conditioned on either Z stabilizer registering a -1 value. If there are bit-flip errors on any two data qubits, this decoding scheme may fail by applying an erroneous correction to the data. Lastly, the error-detection panel corresponds to when both stabilizers register a $+1$ value, indicating no error. Experimental shots which do not satisfy this requirement are discarded. As a result, this decoding scheme only fails with three or more bit-flips, at the expense of discarding data.

Using the FT circuit (blue) and performing error correction in post-processing we prepare $|0\rangle_L, |1\rangle_L, |+\rangle_L$, and $|-\rangle_L$ states with respective errors $0.21(4)\%, 0.39(5)\%, 0.71(7)\%$, and $1.04(9)\%$. We note that the state preparation and measurement error for a single physical qubit in the $|1\rangle$ state is $0.71(4)\%$, limited by bright-to-dark pumping errors during detection. This is one context in which the logical qubit clearly outperforms our physical qubit. For the non-FT circuit (red) the respective errors are $0.93(8)\%, 1.05(9)\%, 3.7(2)\%$, and $3.8(2)\%$, demonstrating the clear benefits of fault-tolerance, but also displaying high-fidelity creation of a globally entangled 9-qubit state. This non-FT preparation circuit can be used to create the $|H\rangle_L = e^{-i\pi\bar{Y}/8}|0\rangle_L$ magic state, which can be used for distillation[32,44].

Phase flip errors can be measured by preparing $|+\rangle_L$, waiting some time, applying varying $Z(\phi)$ gates to every data qubit, and then performing a logical gate $Y_L(-\pi/2)$, as in a typical Ramsey experiment. At each wait time, the Ramsey fringe is fit to extract the contrast (see Supplementary Information) and the resulting contrast as a function of time is fit to a decaying exponential Ae^{-t/T_2^*}. The results of this logical T_2^* experiment are presented in Fig. 2(c). For the raw, error correction, and error detection decoding schemes, we measure a T_2^* of 27(2) ms, 78(9) ms, and 300(90) ms. The measured T_2^* of each independent GHZ state in the logical qubit is almost entirely explained by the measured $T_2^* = 0.6(1)$ s of the individual physical qubits (see Supplementary Information). We attribute the physical qubit decoherence primarily to control noise. With microwave pulses, which are not sensitive to optical beam path fluctuations, and using dynamical decoupling to suppress magnetic field inhomogeneity, we measure $T_2 = 2.84(16)$ s for the physical qubits. We note that with further improvements to the experimental apparatus, $T_2 > 1$ hour can be achieved[45].

Stabilizer Measurements

Having demonstrated high-fidelity preparation of logical states, we proceed to demonstrate FT syndrome extraction on the encoded qubit[39]. In Fig. 3 we show the results of using direct measurement of the stabilizers with four additional ancilla qubits. First, the state is fault-tolerantly encoded into the $|0\rangle_L$ state. Then, an artificial error is applied to a data qubit.

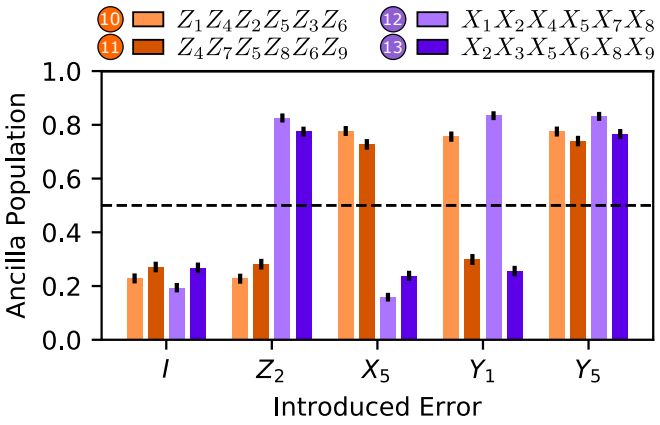

Figure 3. **Detection of arbitrary single-qubit errors.** After encoding $|0\rangle_L$, different Pauli errors are purposely introduced on a selected data qubit in the code. To detect the error, each stabilizer value is mapped onto the state of the corresponding ancilla qubit. The ideal ancilla population is $0/1$ depending on whether an error *did not*/*did* occur on the stabilizer block. The colored bars correspond to the measured population of the different ancilla qubits. The error bars indicate the 95% binomial proportion confidence interval.

Finally, the X and then Z stabilizers are measured sequentially. If no error has occurred, all four stabilizers commute with the logical qubit state and the ancilla qubits should remain in the $|0\rangle$ state. Conversely, if an error did occur on a data qubit, the stabilizers that do not commute with that error flip the state of the ancilla to $|1\rangle$. For example, a Pauli Y error on data qubit 1 commutes with neither the X nor Z stabilizers that measure it, resulting in a flip of ancilla qubits 10 and 12. This confirms our ability to identify arbitrary single qubit errors along both X and Z axes using the stabilizer outcomes. The data presented in Fig. 3 are a representative sample of selected errors; the full data set is available in Supplementary Information. Averaged over all the artificial errors, the measured ancilla qubits 12, 13, 10, and 11 (in order of measurement) differ from the expected value by $17.9(3)\%$, $24.8(3)\%$, $24.4(3)\%$, and $29.8(6)\%$, respectively. This error is due to state preparation, dephasing and cumulative errors introduced by preceding stabilizer measurements. These results are well explained by the $3.8(2)\%$ raw $|0\rangle_L$ encoding error, $6.9(5)\%$ error per X stabilizer, $6.4(7)\%$ error per Z stabilizer, and a fixed $7.2(5)\%$ Z-type error on the logical qubit that is consistent with the expected raw T_2^*-decay over the 3 ms time required to measure X stabilizers, as shown in Fig. 2(c).

Logical Gates

Logical operations are implemented on a quantum error correcting code through circuits that manipulate the entangled state of the physical qubits. We implement a $Y_L(\theta)$ rotation, which can only be implemented transversally for a discrete set of angles[46]. In the case of Bacon-Shor, the smallest transversal $Y_L(\theta)$ rotation we can create is $Y_L(\pi/2)$, which

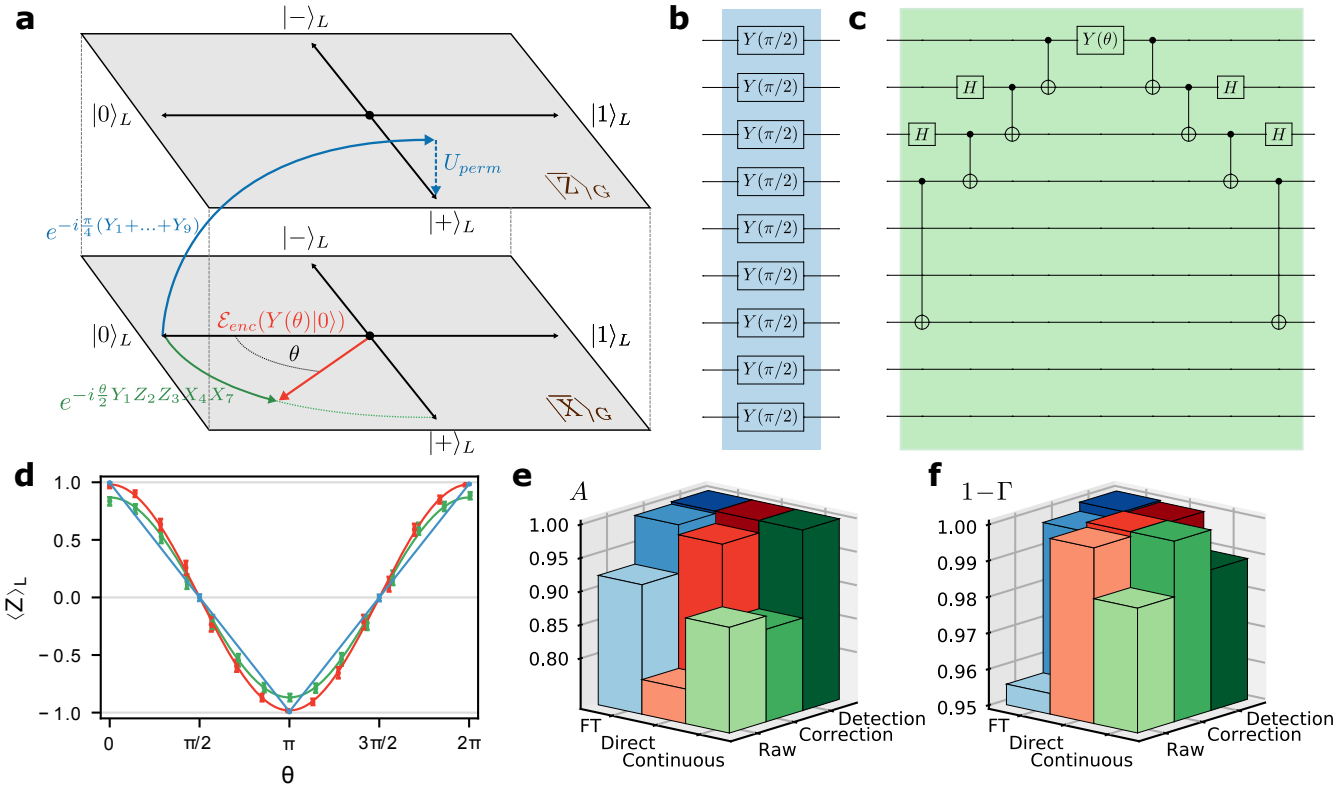

Figure 4. **Manipulating logical states. a,** A schematic depicting different logical operations. A FT discrete logical rotation (blue) operating on $|0\rangle_L$ is a transversal operation, $Y_L(\pi/2) = Y(\pi/2)^{\otimes 9}$, that leaves the code subspace (gray planes) and returns via a permutation of qubit labeling (U_{perm}). A non-FT continuous logical rotation (green) operating on $|0\rangle_L$ is a 5-qubit entangling operation, $Y_L(\theta) = Y_1Z_2Z_3X_4X_7(\theta)$, that rotates through the code subspace. At $\theta = \pi/2$, these gates are equivalent up to a gauge operation. Finally, one can directly prepare $\cos(\frac{\theta}{2})|0\rangle_L + \sin(\frac{\theta}{2})|1\rangle_L$ (red) via a single qubit rotation on one qubit in the code (circuit shown in Fig 2a). **b,** The circuit for the FT gate shown by the blue curve. **c,** The circuit for the non-FT gate capable of creating any state along the green curve. **d,** Experimental results comparing the three logical operations after applying error correction. The expectation value of the logical Z operator is fit to a decaying sinusoid $\langle Z\rangle_L = A\cos(\theta)e^{-\Gamma\theta/\frac{\pi}{2}}$. The error bars are 95% confidence intervals from the binomial distribution. **e,** The amplitude parameters (A) from the fits to the data. This represents the overhead error associated with the operation at $\theta = 0$. **f,** The decay parameters (Γ) from the fits to the data. This represents the error associated with the operation per $\pi/2$ angle.

is generated by applying a physical $Y(\pi/2)$ to each data qubit, followed by relabeling the data qubit indices in post-processing (blue, Fig. 4a-b). We compare the performance of this FT rotation with a non-FT circuit which implements $Y_L(\theta) = Y_1Z_2Z_3X_4X_7(\theta)$ (green, Fig. 4a). These rotations are equivalent for $\theta = \pi/2$ on the logical qubit, but differ in their operation on the gauge qubits. This non-FT gate commutes with both the stabilizers of the code and the gauge operators; if this were not the case, then small angle rotations would entangle the logical qubit with the gauge qubits, spoiling the error protection. The non-FT gate (Fig. 4c) also entangles the separate GHZ states, and so the failure of a single operation in the circuit can lead to the failure of the logical qubit. Finally, we compare both these rotations to the non-FT direct encoding circuit in Fig. 2a (red, Fig. 4a).

The results of these different gate operations, after error-correction, are shown in Fig. 4d. The Rabi curves are fit to a decaying sinusoid $\langle Z\rangle_L = A\cos(\theta)e^{-\Gamma\theta/\frac{\pi}{2}}$ and the fit parameters are presented in Fig. 4e-f. The gate error, corresponding to Γ, for the FT gate is 0.3(1)% after error-correction. The error rate measured here for the FT gate explains the additional error present for the $|+/-\rangle_L$ states in Fig. 2b, which requires two additional $Y_L(\pi/2)$ gates for state preparation and measurement. The rest of the numerical values are tabulated in the Supplementary Information. Derived from these parameters, the error at $\theta = \pi$, the maximum gate angle required with optimized circuit compilation, is 0.45%, 6.59%, and 0.93% for FT gates, non-FT continuous rotations, and non-FT direct encoding, respectively. We note that the loss of contrast when applying error correction to the continuous rotation is indicative of a high proportion of weight-2 errors relative to weight-1 errors. From this data, there is a clear benefit to fault-tolerance, as expected. The results also indicate that for preparation of arbitrary states in the xz-plane, direct encoding is preferable to a continuous rotation.

Outlook

In this work, we have demonstrated high-fidelity operation of a logical qubit capable of correcting any single-qubit errors. The next milestone is to demonstrate a tranversal CNOT logical gate that outperforms the physical two-qubit gate, which is the limiting operation in ion systems. This experiment should be possible in the current system given that two-qubit gates on 23 data qubits have recently been demonstrated [47]. Additionally, multiple rounds of error-correction can be achieved by breaking the ion chain to perform mid-circuit detection[48]. This shuttling will likely require sympathetic cooling schemes, which have been previously demonstrated[49,50] and can be readily implemented in this system[47].

Acknowledgments

We acknowledge fruitful discussions with N. Linke and the contributions of J. Mizrahi, K. Hudek, J. Amini, K. Beck, and M. Goldman to the experimental setup. This work is supported by the ARO through the IARPA LogiQ program, the NSF STAQ Program, the AFOSR MURIs on Dissipation Engineering in Open Quantum Systems and Quantum Interactive Protocols for Quantum Computation, and the ARO MURI on Modular Quantum Circuits. L. Egan and D. M. Debroy are also funded by NSF award DMR-1747426.

Author Contributions

L.E. collected and analyzed the data. L.E., D.M.D., C.N., and M.N., wrote the manuscript and designed figures. M.C. and C.M. led construction of the experimental apparatus with contributions from L.E., C.N., A.R., D.Z., and D.B. Theory support was provided by D.M.D., M.N., M.L., and K.R.B.. C.M. and K.R.B. supervised the project. All authors discussed results and contributed to the manuscript.

Methods

Experimental implementation

We trap $^{171}Yb^{+}$ in a microfabricated-chip ion trap (High Optical Access 2.1.1 from Sandia National Labs) driven by an RF voltage at a frequency of 36.06 MHz. We define the x-axis along the trap axis, with the z-axis perpendicular to the chip surface. A magnetic field of 5.183 G along the z-axis defines the atomic quantization axis. The individually-addressing (global) Raman beam is oriented along the $z(y)$-axis of the trap, so that the Raman process transfers momentum to the ions along the $\hat{y}-\hat{z}$ direction. We selectively couple light to the lower-frequency set of radial modes by tilting the trap principal axes using a static electric yz quadrupole. We use quadratic and quartic axial potentials to minimize the spacing inhomogeneity for the middle N-2 ions. In the 15-ion chain, the longest wavelength (in-phase) mode along each trap axis is $(\nu_x, \nu_{y-z}, \nu_{y+z}) = (0.193, 3.077, 3.234)$ MHz.

An imaging objective with numerical aperture 0.63 (Photon Gear, Inc.) is used to focus each of the 32 individual beams to a waist of 0.85 μm, spaced by 4.43 μm at the ions. The mode-locked 355 nm laser (Coherent Paladin 355-4000) used to drive Raman transitions has been modified to tune the repetition rate of the laser so as to null the 4-photon cross-beam Stark shift. Typical spin-flip Rabi frequencies achieved in our system are 500 kHz. The maximum crosstalk on nearby ions is 2.5% of the Rabi frequency of the addressed ion.

Before each experiment, the ions are cooled to near the motional ground state through a combination of Doppler cooling and Raman sideband cooling. After the circuit, resonant 369 nm light on the $^2S_{1/2} \rightarrow {}^2P_{1/2}$ cycling transition is used to perform state detection. Scattered light is collected through the 0.63 NA objective and imaged with magnification of 28 onto a multi-mode (100 μm core) fiber array that is broken out into individual photo-multiplier tubes (Hamamatsu H10682). About 1% of the total light is detected as counts. Dark/bright states are mapped to $|0\rangle/|1\rangle$ states by setting a threshold at > 1 photon detected within a detection window (typically 100 μs). State preparation and detection errors are $0.22(2)\%$ and $0.71(4)\%$ for $|0\rangle$ and $|1\rangle$. Detection crosstalk onto neighboring channels is $0.3(2)\%$; see Supplementary Information for detailed error budget.

The entire experiment is controlled by an FPGA (Xilinx) programmed via the ARTIQ software. RF gate waveforms are generated by a 4-channel AWG (Keysight M3202A), one of which drives the global beam, and two of which are routed through a custom switch network onto any of the 15 middle channels of the individual beam AOM at each timestep in the circuit.

Native ion-trap single-qubit gates

The native physical single-qubit gate available to our system is a single qubit rotation about a vector in the xy-plane, $R(\theta,\phi)$ where θ is the angle of rotation and ϕ is the angle between the rotation axis and the x-axis. In this notation, $RX(\theta) = R(\theta,0)$ and $RY(\theta) = R(\theta,\pi/2)$. Additionally, we use compound SK1 pulses to suppress angle and cross-talk errors[51]. The SK1 pulses are shaped with a smooth Gaussian amplitude envelope to avoid frequency content that may excite axial motion due to light-induced prompt charge effects from partially exposed semiconductor in the chip trap. Due to hardware limitations, single-qubits gates are run sequentially. We implement virtual $RZ(\theta)$ gates via a software advance of the local oscillator phase, tracked for each individual ion. Before each circuit is run, we calibrate the amplitude of an $RX(\theta)$ on each qubit in the chain. We achieve single-qubit native gate error rates of $1.8(3)\times10^{-4}$ on a 15-ion chain as measured by randomized benchmarking (see Supplementary Information).

Native ion-trap two-qubit gates

The native two-qubit operation is the $XX(\theta)$ Ising gate, implemented via a Mølmer-Sørensen interaction[52]. CNOT gates can be constructed from an $XX(\pi/4)$ gate and additional single qubit gates[53]. Offline, we calculate laser pulse solutions for XX gates to disentangle the motional modes using amplitude-modulated waveforms[34] discretized into 16 segments with

linear interpolation between segments to avoid undesirable excitation of the axial motion. In an equispaced chain of 15 ions, we observe that the middle 11 radial motional modes are also roughly equispaced. The laser detuning from motional modes is constant across the waveform and is chosen to sit approximately halfway between two adjacent modes, which leads to particularly simple laser waveforms to eliminate qubit-motion entanglement at the end of the gate. The gate frequency for a particular gate pair is optimized to minimize the required laser power, minimize sensitivity to mode-frequency errors of < 1 kHz, and to avoid coupling to modes with low spatial frequencies that are subject to heating . Gate durations are 225 μs. To avoid unwanted couplings, we run two-qubit gates sequentially. Before a batch of circuits is run, we calibrate the amplitude, common phase and differential phase of each gate in the circuit. We achieve between 98.5 and 99.3% fidelity on a typical gate, measured by parity fringes after a varying odd number of successive of non-echoed and echoed XX gates (see Supplementary Information).

Crosstalk Detection

When available, idle qubits in a circuit are used to detect potential two-qubit gate crosstalk errors. Any experimental shots where an idle qubit is measured in the $|1\rangle$ state are discarded in post-processing. On average, $< 4\%$ of the total data is discarded using this method.

Supplementary Information

State preparation and measurement errors

We characterize the state preparation and measurement (SPAM) errors using the following method. We load a single ion and prepare it in the $|0\rangle$ state using optical pumping, from which we may also apply an SK1[51] π-rotation to prepare $|1\rangle$. To measure the qubit state, 369 nm light that is resonant with the $\{^2S_{1/2}, F=1\} \longleftrightarrow \{^2P_{1/2}, F=0\}$ transition is shined onto the ion and the scattered photons are detected using our array of PMTs. We determine that the ion is bright (dark) when we detect > 1 (≤ 1) photons within a 100 μs window. When the ion is prepared in $|1\rangle$ (the bright state), we measure a SPAM error of $0.71(4)\%$, while the SPAM error for $|0\rangle$ (the dark state) is $0.22(2)\%$, making the average single-qubit SPAM error $0.46(2)\%$. Table S1 describes the SPAM error budget, derived either from separate measurements or by fitting Poisson curves to the histogram of photon count event frequency. The measured average detection cross-talk to neighboring PMTs when the target ion is bright is $0.3(2)\%$.

SK1 pulse, 1-state and 0-state error	**Error budget**
SPAM error on bright ion $\equiv |1\rangle$	0.71%
Bright to dark pumping	0.55%
Thresholding error	0.12%
Preparation error (1-qubit randomized benchmarking)	0.03%
SPAM error on dark ion $\equiv |0\rangle$	0.22%
Dark to bright pumping	0.13%
Preparation error - incomplete pumping	0.02%
Background dark counts (measured with no ion qubit)	0.07%
Detection cross-talk error (averaged across neighboring PMTs to bright ion)	0.34%

Table S1. State preparation and measurement error budget for a single ion in our system.

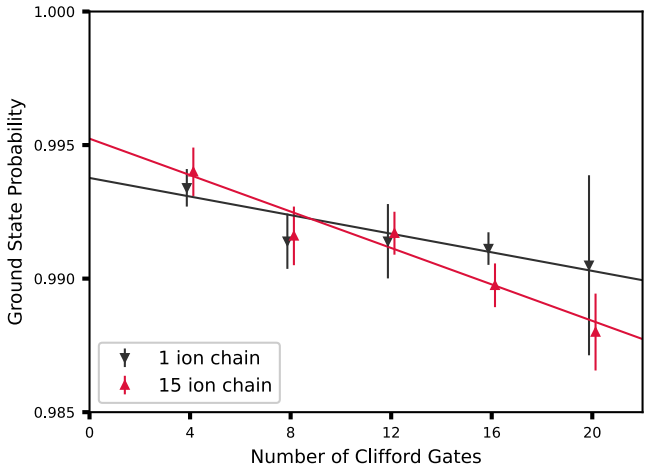

Figure S1. **Randomized benchmarking of single-qubit gates.** The probability to measure a single ion in the ground state after a variable number of Clifford gates in a randomized benchmarking sequence. The slope of the line indicates the per-Clifford fidelity, while the y-intercept indicates the SPAM error. Fit function for 1 ion chain is $0.9938(8) - N * 1.7(7) \times 10^{-4}$, and for 15 ion chain $0.995(1) - N * 3.4(8) \times 10^{-4}$. Error bars shown are the standard error of the mean.

Single qubit gate benchmarking

The reported single qubit gate fidelity was measured using single qubit randomized benchmarking[54], using a sequence of up to 20 random Clifford gates, which were decomposed into our native rotation gates and implemented using SK1 composite pulses. Each random sequence is followed by its inverse in order to, in principle, echo out the gates completely and return the qubit to the initial $|0\rangle$ state. The degree to which the qubit does not return to the initial state quantifies the infidelity of the circuit. The measured occupation of the $|0\rangle$ ground state as a function of the number of the applied Clifford gates is shown in Fig. S1. This benchmarking procedure is performed on a single ion, as well as on an individual qubit in a chain of 15 ions, so as to detect any

adverse affects arising from an increase in the system size. The fitted slope of the occupation of the $|0\rangle$ state as a function of the number of the applied Clifford gates indicates a per-Clifford error of $3.4(8)\times10^{-4}$ on the 15-ion chain, corresponding to an error of $1.8(3)\times10^{-4}$ per native Pauli gate[55]. The offset in the fit is consistent with SPAM errors.

Two-qubit gate performance

Two qubit gate performance was estimated using the results of two gate sequences. On gates with low crosstalk (details in the following section), we anticipate the dominant error in the $XX(\pi/4)$ gate will be an over or under-rotation error by a small angle ε resulting in $XX(\pi/4+\varepsilon)$. A sequence of N successive applications of the gate will then will result in an accumulated error of $N\varepsilon$. If the phase of the gate is flipped with each successive application, $XX(\pi/4+\varepsilon)XX(-\pi/4-\varepsilon)...$, then this particular error is suppressed to the extent that it is stable between applications. We also note that the echo sequence will also suppress other forms of coherent errors, such as gate crosstalk. By comparing the resultant error of the two sequences (with and without the echo), we can estimate the true fidelity of a single gate to be between the two. The fidelity is determined by the parity fringe method[56]. We increase the number of XX gates in the sequence, and take the slope of the fidelity to be the error per gate. The results of this estimation are shown for a single gate between ions 2 and 3 in the chain of 15 in Fig. S2. In other gates, a decrease in fidelity relative to this estimate is primarily due to effects of crosstalk.

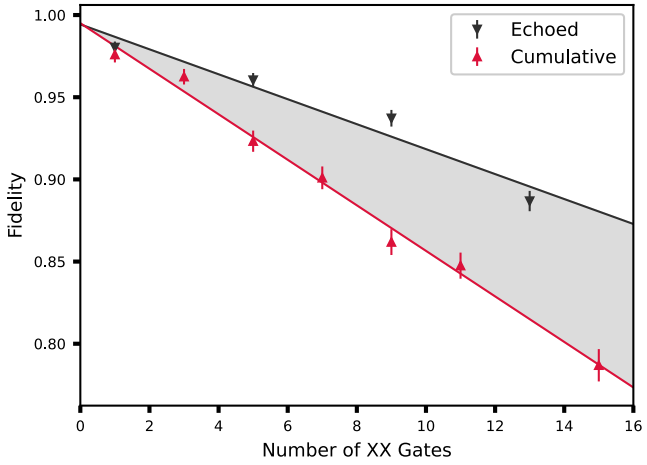

Figure S2. **Two-qubit gate performance.** The cumulative sequence (red) $XX(\pi/4+\varepsilon)XX(\pi/4+\varepsilon)...$ results in a total error of $N\varepsilon$, and the echoed sequence (black) $XX(\pi/4+\varepsilon)XX(-\pi/4-\varepsilon)...$ cancels this particular type of over/under-rotation error. The true gate fidelity lies somewhere in the shaded region. The slope of the two lines determines the error per gate, or the estimated fidelity to be in the range of $98.5-99.3\%$. Errors are plotted for 95% confidence intervals.

Circuit optimization for crosstalk

There are several factors to consider when mapping the Bacon-Shor code onto a chain of 15 ions, as shown in Fig. 1 of the main text. In general, ion chains feature all-to-all two-qubit gate connectivity; however, some gates require more optical power than others to achieve maximal entanglement. Considering errors that scale with intensity, such as crosstalk, then gate fidelity is expected to decrease with increasing power requirements. These differences in power requirements can be understood by examining the mode participation symmetries in the chain. For example, ion 8, the center ion, requires high power in nearly all of its gates because it only participates in the even spatial modes (i.e., $b_{8,2n}=0$, $n=1,2,...,7$ where $b_{i,1}$ is mode-participation factor of the highest-frequency in-phase radial mode for ion i). So on average, for a fixed gate frequency, the modes that drive entanglement are further detuned from the gate. We note that this is unique to our choice of amplitude modulated (AM) gates with a fixed frequency; phase/frequency-modulated (PM/FM) gates or multi-tone gates may have different chain symmetry considerations.

In Fig. S3, we present the power requirements for the gates in our system. We first optimize the frequency of each gate across the mode spectrum to find pulse solutions that are robust to mode errors of <1 kHz. Once the frequency is fixed for each gate, we calculate the root-mean-square (RMS) Rabi frequency (Ω_{rms}) of the AM waveform for *each* red/blue sideband when brought into resonance with the carrier transition. In our system, we use equal Rabi frequencies to drive both ions i,j in the gate ($\Omega_{i,rms}=\Omega_{j,rms}$), although this need not be case. The Lamb-Dicke factor ($\eta\approx0.08$) converts carrier Rabi frequency to sideband frequency and this factor is normalized by the gate duration ($\tau_{gate}=225\mu$ s). Using Fig. S3 as a cost matrix, we manually optimize the mapping so that the required gates in the circuit minimize the total cost. In general, we observe that each half of the chain has strong coupling to itself, and the two halves of the chain couple well to each other as long as symmetry

of the chain is obeyed (e.g., gates where the ions are with both odd or both even integer offsets from the center of the chain couple well, but mixed even and odd integer offsets do not). We further note that when considering the full stabilizer circuit, it is preferable to use ion 8 as a data qubit (maximum 4 gates) than as an ancilla qubit (6 gates). With these considerations, we arrived at the ion→qubit mapping displayed in Fig. 1 of the main text.

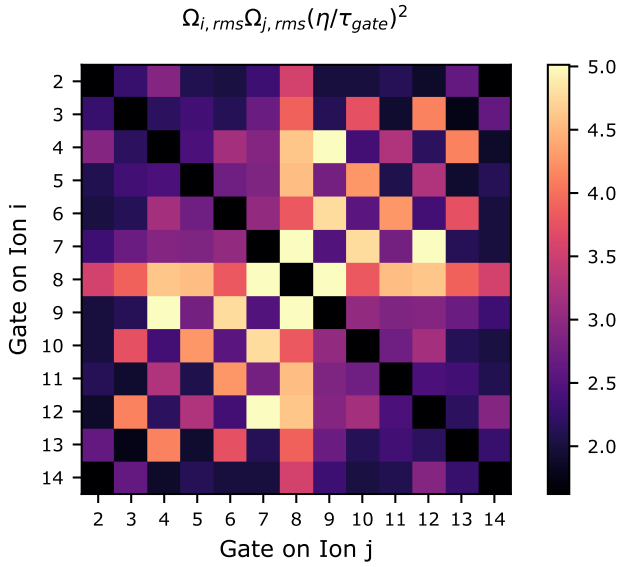

Figure S3. **Power requirements for XX gates in our system.** For an XX gate on ions i, j, the (RMS) Rabi frequency (Ω_{rms}) of the AM waveform for *each* red/blue sideband when brought into resonance with the carrier transition is normalized by the Lamb-Dicke factor ($\eta \approx 0.08$) and the gate duration ($\tau_{gate} = 225\mu$s). Gate power is used as a proxy for crosstalk, and the ion→qubit mapping is chosen to minimize the cost matrix.

Logical T_2^* fits

The Ramsey fringe amplitudes shown in Fig. 2c are calculated by fitting a curve to a logical Ramsey experiment at each wait time. The data is taken by preparing a $|+\rangle_L$ state as shown in Eq. 1, waiting some amount of time t, applying varying $RZ(\theta)$ gates to every qubit and then measuring in the logical X basis via a transversal $RY_L(-\pi/2)$. Here we will explain the theoretical fits used for raw, corrected, and detected data processing techniques.

Firstly, as shown in Eq. 1, the logical $|+\rangle$ state we use is composed of three GHZ states $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$. Due to the structure of these states, if a $RZ(\theta)$ gate is applied to each qubit, the three gates will coherently combine, and the end result will be the same as if a $RZ(3\theta)$ gate had been applied on any single qubit. By considering this simplification we can reduce the number of error cases we must consider.

For the 'raw' processing case, any Z error flips the logical output. As a result the cases where 1 or 3 errors occur lead to $|-\rangle_L$ states, while while cases with 0 or 2 errors lead to $|+\rangle_L$. Consequently the expectation value of X_L can be thought of as the squared amplitude of cases which lead to $|+\rangle_L$, subtracted by the squared amplitude of cases which result in $|-\rangle_L$. This results in a curve

$$\langle X_L\rangle = \cos(3\theta/2)^6 - 3\cos(3\theta/2)^4\sin(3\theta/2)^2 + 3\cos(3\theta/2)^2\sin(3\theta/2)^4 - \sin(3\theta/2)^2 = \cos(3\theta)^3.$$

In the 'corrected' processing case, the state can tolerate a single error without having its logical information corrupted. As a result error cases with 0 or 1 errors lead to $|+\rangle_L$, while 2 or 3 lead to $|-\rangle_L$. This results in the curve

$$\langle X_L\rangle = \cos(3\theta/2)^6 + 3\cos(3\theta/2)^4\sin(3\theta/2)^2 - 3\cos(3\theta/2)^2\sin(3\theta/2)^4 - \sin(3\theta/2)^2.$$

Lastly the 'detected' processing method is slightly more complex, as postselection means we must renormalize the expectation value. The case with 0 errors leads to $|+\rangle_L$, while the case with 3 errors leads to $|-\rangle_L$. Cases with 1 or 2 errors must set off at least one stabilizer, and as a result those runs will be removed from the dataset. As a result the probabilities must be

renormalized, leading to the curve

$$\langle X_L \rangle = \frac{\cos(3\theta/2)^6 - \sin(3\theta/2)^6}{\cos(3\theta/2)^6 + \sin(3\theta/2)}.$$

In an experiment there will also be imperfections in the states due to errors beyond T_2^* dephasing, which we model be a simple depolarization of each GHZ state with strength p. This corresponds to taking a state $|+\rangle_{GHZ} \to (1-p)|+\rangle\langle+|_{GHZ} + \frac{p}{2}(|+\rangle\langle+|_{GHZ} + |-\rangle\langle-|_{GHZ})$, where the second term, equal to the maximally mixed state on the space spanned by $|000\rangle$ and $|111\rangle$, has an expectation value of 0.

In the raw case, any depolarization of the GHZ states will lead to the expectation value going to zero, and as a result the only non-zero expectation values come about when no depolarization occurs. As a result the overall fringe pattern is simply scaled by a factor of $(1-p)^3$:

$$\langle X_L \rangle = A(1-p)^3 cos(3\theta)^3.$$

In the correction case the stabilizers are able to identify and correct a single depolarization error. This leads to different cases for when depolarization occurs and when they do not, which when collated lead to:

$$\begin{aligned}\langle X_L \rangle = &\left[(1-p)^3 + 3(1-p)^2 p + \frac{3p^2(1-p)}{2}\right](\cos(3\theta/2)^6 - \sin(3\theta/2)^6) \\ &+ \left[3(1-p)^3 + 3(1-p)^2 p + \frac{3p^2(1-p)}{2}\right](\sin(3\theta/2)^2\cos(3\theta/2)^4 - \sin(3\theta/2)^4\cos(3\theta/2)^2).\end{aligned}$$

The most complex case is the detection case. Individual depolarizations each contribute a $\frac{1}{2}$ chance of setting off a stabilizer, and when they do not the coherent rotations on the other qubits produce similar behaviors to the ideal detection case, but only on the non-depolarized qubits. This leads to the equation:

$$\begin{aligned}\langle X_L \rangle = A\Bigg[&(1-p)^3 \left(\frac{\cos(3\theta/2)^6 - \sin(3\theta/2)^6}{\cos(3\theta/2)^6 + \sin(3\theta/2)^6}\right) \\ &+ \frac{3p(1-p)^2}{2}\left(\frac{\cos(3\theta/2)^4 - \sin(3\theta/2)^4}{\cos(3\theta/2)^4 + \sin(3\theta/2)^4}\right) \\ &+ \frac{3p^2(1-p)}{4}\left(\frac{\cos(3\theta/2)^2 - \sin(3\theta/2)^2}{\cos(3\theta/2)^2 + \sin(3\theta/2)^2}\right)\Bigg].\end{aligned}$$

These models well describe the experimental data, as shown in Fig. S4.

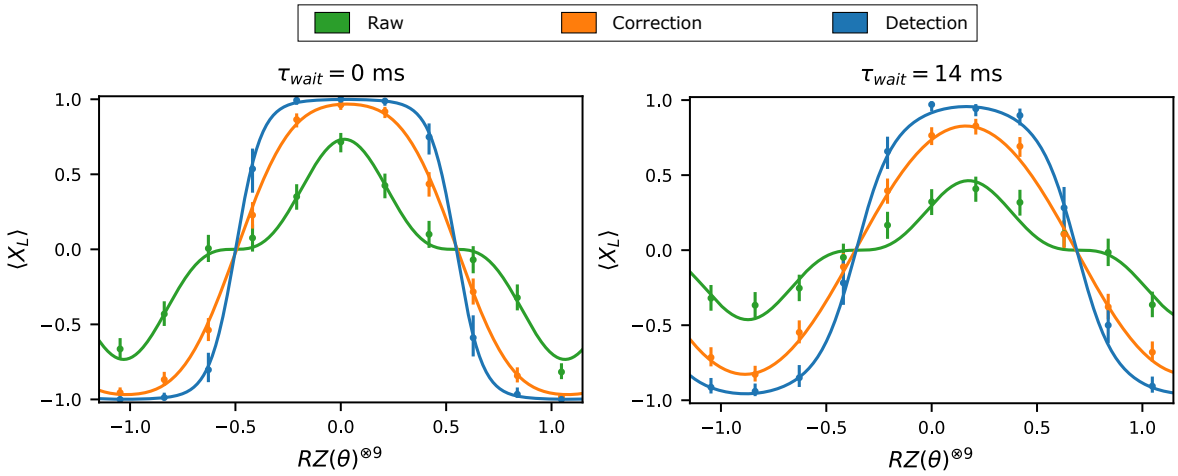

Figure S4. **Examples of T_2^* fits.** These plots exemplify logical Ramsey fringe fitting at two different wait times, 0ms (left) and 14ms (right). At short wait times (left), the data shows characteristics of error correction, such as the flattened top for error-detection. At longer wait times (right), the amplitude of each curve decreases due to T_2^*, but also the flat features of the curves blur due to GHZ depolarization. In both cases, the error model well matches the experimental data. Error bars are the 95% binomial proportion confidence interval

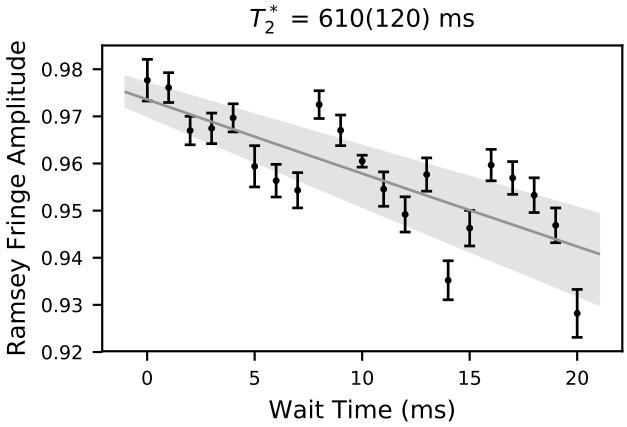

Figure S5. **Raman T_2^* for a physical qubit in a 15-ion chain.** Ramsey fringes with variable wait times are fit to a sinusoud to extract the contrast. For each data point, the error bars represent the 95% confidence interval from a maximal likelihood estimation fit of a sinuosoid. The contrast is fit to a decaying exponential $Ae^{-\tau_{\text{wait}}/T_2^*}$, with fit value $T_2^* = 610(120)$ ms. The shaded region indicates the 1σ uncertainty in the exponential least-squares fit.

Physical T_2^*

To understand the phase flip errors in the logical qubit, we measure the T_2^* of the physical qubits in a chain of 15 ions. This is accomplished via a laser Raman Ramsey sequence on the center ion, $RY(\pi/2) - \tau_{\text{wait}} - RZ(\theta) - RY(-\pi/2)$, with no echoes. At each wait time τ_{wait}, the phase θ is swept, and the resulting data is fit to a sinusoid to extract the contrast. The Ramsey contrast is fit to a decaying exponential $Ae^{-\tau_{\text{wait}}/T_2^*}$ to extract T_2^*. The results of this experiment are shown in Fig. S5. We find $T_2^* = 610(120)$ ms for a physical qubit in a chain of 15 ions. We attribute the physical qubit decoherence primarily to control noise, rather than to fundamental qubit decoherence. In particular, we note that there are features of revivals at ≈ 8 ms and 16 ms, corresponding to noise at ≈ 125 Hz. We assign this to mechanical fluctuations (e.g., fans) that shift the standing wave of the optical Raman beams relative to the ions. This effect can be mitigated by switching to a "phase-insensitive" configuration[57].

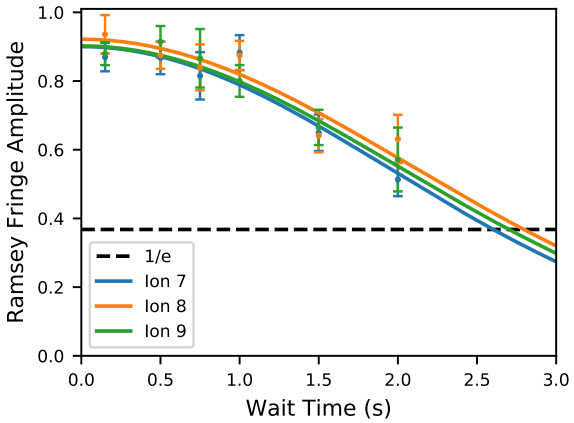

Figure S6. **Microwave T_2 for a physical qubit in a 15-ion chain.** An $(XY)^N$ dynamical decoupling sequence is used to suppress magnetic field noise. Ramsey fringes with variable wait times are fit to a sinusoud to extract the contrast. For each data point, the error bars represent the 1σ confidence interval from a least-squares fit of a sinuosoid. The contrast is fit to a Gaussian decay $Ae^{-(\tau_{\text{wait}}/T_2)^2}$, with average fit value $T_2 = 2.84(16)$ s.

To investigate the degree to which control errors impact our physical T_2^* qubit decoherence, we perform microwave Ramsey experiments, which are not sensitive to optical beam path fluctuations. Additionally, we suppress magnetic field inhomogeneity using a dynamical decoupling technique that applies π-pulses with alternating 90° phase offsets, commonly known as an $(XY)^N$ pulse sequence[58], to periodically refocus the qubit spin. Due to a strong microwave drive field gradient along the ion chain, the π-pulse times are only well calibrated for three neighboring ions in the chain, which we center on the middle ion (8) in the chain. At each wait time, the phase of the fringe is swept, and the resulting data is fit to a sinusoid to extract the contrast. The

resulting data is shown in Fig. S6. We observe that the resulting decay is better fit to a Gaussian ($Ae^{-(\tau_{\text{wait}}/T_2)^2}$), compared to an exponential decay, with average $T_2 = 2.84(16)$ s. The coherence time of this echo experiment is limited by residual magnetic field noise, which can be improved by operating our qubit in a lower bias-field or by using magnetic shielding. We note that $T_2 > 1$ hour has been achieved in $^{171}\text{Yb}^{+}$[45].

The GHZ states that make up the Bacon-Shor code ($|000\rangle + |111\rangle$) are three times as sensitive to phase noise as our physical qubit. To understand the implication of the Raman T_2^* on the logical T_2^*, we run numerical simulations to extrapolate the measured phase noise to a GHZ state. We assume that the Pauli-Z noise in the middle of the Ramsey sequence is Gaussian distributed with some width Δ_Z. Using the fit from Fig. S5, we can numerically solve for the width of the noise spectrum Δ_Z. Once this value is found, we re-run the simulation with that noise spectrum on a three-qubit GHZ state to extract the predicted contrast. In Fig. S7, we compare this predicted value to the three individual GHZ states measured in the logical qubit experiment. We conclude that almost all the dephasing in the logical qubit that we observe is explained by the observed T_2^*- decay in the physical qubit. We note that this is the same experimental data presented in Fig. 2c of the main text, just post-processed to analyze individual GHZ states rather than to perform error-correction.

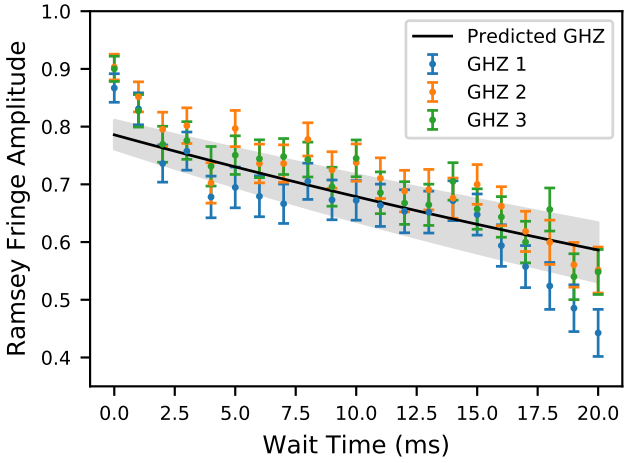

Figure S7. T_2^* **for GHZ states.** Ramsey fringes for each independent GHZ state are analyzed from the logical T_2^* experiment in Fig. 2c of the main text. For each data point, the error bars represent the 95% confidence interval from a maximal likelihood estimation fit of a sinuosoid. For the theoretical prediction, we apply the fitted dephasing noise in Fig. S5 to a numerical simulation of a three-qubit GHZ state. The shaded region indicates the 1σ uncertainty in the exponential least-squares fit in Fig. S5, propagated through the numerical simulation.

Extended stabilizer results

In Fig. 3 of the main text, we presented a representative sample of artificially introduced errors and the corresponding ancilla qubit populations. Here in Fig. S8, we present a full set of errors that produce all of the possible ancilla qubit output bit strings.

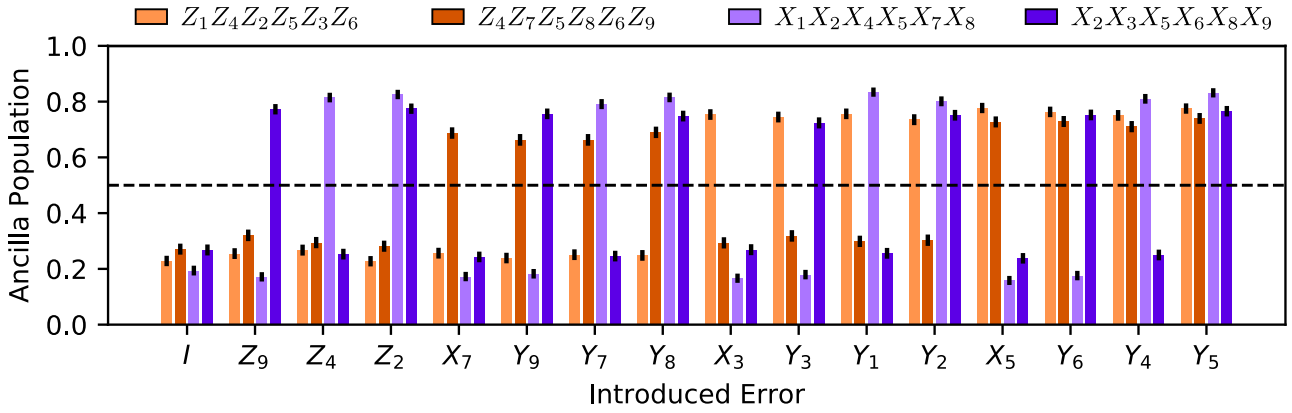

Figure S8. **Full stabilizer measurement results.** A complete set of errors are introduced to create all possible output bit strings of the ancilla qubits. Error bars are the 95% binomial proportion confidence interval.

For a given error, each stabilizer measurement yields a deterministic eigenvalue measurement (e.g., $\{+1,-1,-1,+1)\}$ that is mapped to the ancilla qubit state (e.g., $\{0,1,1,0\}$). Defining the error as the difference between the expected ancilla bit string and the measured populations, and averaging across all the artificial errors, we obtain the following total error for each stabilizer measurement:

Stabilizer	**Total Error (ε_{S_i})**
$S_1 = Z_1Z_4Z_2Z_5Z_3Z_6$	$0.244(3)$
$S_2 = Z_4Z_7Z_5Z_8Z_6Z_9$	$0.298(6)$
$S_3 = X_1X_2X_4X_5X_7X_8$	$0.179(3)$
$S_4 = X_2X_3X_5X_6X_8X_9$	$0.248(3)$

In this experiment, the stabilizers are measured in the order S_3, S_4, S_1, S_2. We note from the data presented in Fig. 2 of the main text, the raw encoding of the $|0\rangle_L$ state has a base $\varepsilon_{enc} = 0.038(2)$ error, which we assume is isotropic in the sense that all stabilizer measurements should see the error equally. Additionally, stabilizer measurements will detect errors introduced by itself or previous stabilizer measurements, which we assume to be isotropic as well. The per stabilizer error can be calculated by the differential error between successive stabilizer measurements. We calculate $\varepsilon_Z = 0.064(7)$ per Z-stabilizer (avg. 98.9% gate fidelity) and $\varepsilon_X = 0.069(5)$ per X-stabilizer (avg. 98.8% gate fidelity). Finally, we observe an error offset on the X-stabilizers relative to the Z-stabilizers of $\varepsilon_{T_2^*} = 0.072(5)$, consistent with a Z-type error caused by the logical qubit dephasing (T_2^*) over the wall-clock time it takes to measure the X-stabilizers (≈ 3 ms), as presented in Fig. 2c of the main text. In conclusion, we find that the total stabilizer measurement error for each ancilla qubit is well explained by the following error model:

$$\begin{aligned}
\varepsilon_{S_1} &= \varepsilon_{enc} + 2\varepsilon_X + \varepsilon_Z \\
\varepsilon_{S_2} &= \varepsilon_{enc} + 2\varepsilon_X + 2\varepsilon_Z \\
\varepsilon_{S_3} &= \varepsilon_{enc} + \varepsilon_{T_2^*} + \varepsilon_X \\
\varepsilon_{S_4} &= \varepsilon_{enc} + \varepsilon_{T_2^*} + 2\varepsilon_X
\end{aligned}$$

Bacon-Shor gauge operators

The Bacon-Shor code is an example of a subsystem quantum error correcting code. These codes have additional quantum degrees of freedom which are not protected to the same distance as the logical degree of freedom. In the [[9,1,3]] Bacon-Shor code, there are 4 additional degrees of freedom referred to as *gauge qubits*. One basis for the gauge qubits corresponds to fixing 4 constraints on the eigenvalues of the operators shown in the table below.

X-gauges	Z-gauges
X_1X_2	Z_1Z_4
X_4X_5	Z_2Z_5
X_7X_8	Z_3Z_6
X_2X_3	Z_4Z_7
X_5X_6	Z_5Z_8
X_8X_9	Z_6Z_9

It should be noted that this is not an independent set of operators because the stabilizers of the code, which are products of gauges, already have their eigenvalues fixed to $+1$. As such, if the X-gauges X_1X_2 and X_4X_5 both have eigenvalue $+1$ on a given logical state, then the eigenvalue of X_7X_8 will also be $+1$. We refer to a state in which all $X(Z)$-type gauge operators have eigenvalue $+1$ as the $|\overline{X}\rangle_G(|\overline{Z}\rangle_G)$ gauge. It should be noted that these gauge operators do not commute, so these two gauges are mutually exclusive. When decoding the Bacon-Shor code, we can only identify operators up to a product of gauges. When the correction is applied, we may have inadvertently applied a gauge operator to the logical state. This leaves our logical qubit unaffected, but will alter the state of the gauge qubits.

Logical Pauli operators on a subsystem code decompose as a tensor product of operations on the logical and gauge degrees of freedom. When a logical Pauli operator that acts non-trivially on the gauge subsystem is used to generate a continuous unitary operator, it will entangle the logical and gauge subsystems. As these gauge subsystems are less protected than the logical subsystem, the information will be less protected. Consequently, one must design continuous logical operators around logical Pauli operations that commute with the entire gauge group, ensuring that it acts trivially on the gauge subsystem.

Fit values for logical gate operations

In Table. S2 and S3, we report the numerical values obtained from fitting the logical gate operations as displayed in Fig. 4(e,f) of the main text. Error values are reported as the 1σ from a Gaussian approximation to a maximal likelihood estimation fit.

The Gaussian approximation fails when the fit parameters are at or equal to their extrema (e.g., when $A \simeq 1$), in which case asymmetric error bars are given by the notation $\binom{1\sigma\text{upper}}{1\sigma\text{lower}}$.

Amplitude (A)	Raw	Error Correction	Error Detection
FT Gate	$0.923(6)$	$0.996(1)$	$1.000\binom{0}{3}$
Direct Prep	$0.78(1)$	$0.982(5)$	$0.997(2)$
Continuous Gate	$0.89(1)$	$0.87(1)$	$1.00\binom{0}{1}$

Table S2. Numerical values for the fit parameter A, as presented in Fig. 4e of main text.

Gate Error (Γ)	Raw	Error Correction	Error Detection
FT Gate	$0.044(3)$	$0.0027(7)$	$0.0002(1)$
Direct Prep	$0.001(6)$	$0.000(2)$	$0.001(1)$
Continuous Gate	$0.015(5)$	$0.000(4)$	$0.011(2)$

Table S3. Numerical values for the fit parameter Γ, as presented in Fig. 4f of main text.

References

1. Feynman, R. P. Quantum mechanical computers. *Foundations Phys.* **16**, 507–531 (1986).
2. Abrams, D. S. & Lloyd, S. Simulation of many-body fermi systems on a universal quantum computer. *Phys. Rev. Lett.* **79**, 2586 (1997).
3. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
4. Reiher, M., Wiebe, N., Svore, K. M., Wecker, D. & Troyer, M. Elucidating reaction mechanisms on quantum computers. *Proc. Natl. Acad. Sci.* **114**, 7555–7560 (2017).
5. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.* **41**, 303–332 (1999).
6. Von Burg, V. *et al.* Quantum computing enhanced computational catalysis. Preprint at https://arxiv.org/abs/2007.14460 (2020).
7. Gidney, C. & Ekerå, M. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. Preprint at https://arxiv.org/abs/1905.09749 (2019).
8. Aharonov, D. & Ben-Or, M. Fault-tolerant quantum computation with constant error rate. *SIAM J. on Comput.* (2008).
9. Knill, E., Laflamme, R. & Zurek, W. Threshold accuracy for quantum computation. Preprint at https://arxiv.org/abs/quant-ph/9610011 (1996).
10. Gottesman, D. E. *Stabilizer Codes and Quantum Error Correction*. Ph.D. thesis, California Institute of Technology (1997).
11. Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493 (1995).
12. Knill, E. & Laflamme, R. Theory of quantum error-correcting codes. *Phys. Rev. A* **55**, 900 (1997).
13. Córcoles, A. D. *et al.* Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. *Nat. Commun.* **6**, 1–10 (2015).
14. Takita, M., Cross, A. W., Córcoles, A., Chow, J. M. & Gambetta, J. M. Experimental demonstration of fault-tolerant state preparation with superconducting qubits. *Phys. Rev. Lett.* **119**, 180501 (2017).
15. Linke, N. M. *et al.* Fault-tolerant quantum error detection. *Sci. Adv.* **3**, e1701074 (2017).
16. Harper, R. & Flammia, S. T. Fault-tolerant logical gates in the ibm quantum experience. *Phys. Rev. Lett.* **122**, 080504 (2019).
17. Andersen, C. K. *et al.* Repeated quantum error detection in a surface code. *Nat. Phys.* 1–6 (2020).
18. Cory, D. G. *et al.* Experimental quantum error correction. *Phys. Rev. Lett.* **81**, 2152 (1998).
19. Chiaverini, J. *et al.* Realization of quantum error correction. *Nature* **432**, 602–605 (2004).
20. Schindler, P. *et al.* Experimental repetitive quantum error correction. *Science* **332**, 1059–1061 (2011).

21. Reed, M. D. *et al.* Realization of three-qubit quantum error correction with superconducting circuits. *Nature* **482**, 382–385 (2012).

22. Riste, D. *et al.* Detecting bit-flip errors in a logical qubit using stabilizer measurements. *Nat. Commun.* **6**, 1–6 (2015).

23. Kelly, J. *et al.* State preservation by repetitive error detection in a superconducting quantum circuit. *Nature* **519**, 66–69 (2015).

24. Gong, M. *et al.* Experimental verification of five-qubit quantum error correction with superconducting qubits. Preprint at https://arxiv.org/abs/1907.04507 (2019).

25. Nigg, D. *et al.* Quantum computations on a topologically encoded qubit. *Science* **345**, 302–305 (2014).

26. Luo, Y.-H. *et al.* Quantum teleportation of physical qubits into logical code-spaces. Preprint at https://arxiv.org/abs/2009.06242 (2020).

27. Heeres, R. W. *et al.* Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nat. communications* **8**, 1–7 (2017).

28. Flühmann, C. *et al.* Encoding a qubit in a trapped-ion mechanical oscillator. *Nature* **566**, 513–517 (2019).

29. Ofek, N. *et al.* Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature* **536**, 441–445 (2016).

30. Campagne-Ibarcq, P. *et al.* Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature* **584**, 368–372 (2020).

31. Noh, K. & Chamberland, C. Fault-tolerant bosonic quantum error correction with the surface–gottesman-kitaev-preskill code. *Phys. Rev. A* **101** (2020).

32. Bravyi, S. & Kitaev, A. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A* **71**, 022316 (2005).

33. Maunz, P. L. W. High optical access trap 2.0. *Sandia Natl. Lab. Rep. No. SAND2016-0796R* (2016).

34. Debnath, S. *et al.* Demonstration of a small programmable quantum computer with atomic qubits. *Nature* **563**, 63 (2016).

35. Wright, K. *et al.* Benchmarking an 11-qubit quantum computer. *Nat. communications* **10**, 1–6 (2019).

36. Bacon, D. Operator quantum error-correcting subsystems for self-correcting quantum memories. *Phys. Rev. A* **73**, 012340 (2006).

37. Aliferis, P. & Cross, A. W. Subsystem fault tolerance with the Bacon-Shor code. *Phys. Rev. Lett.* **98**, 220502 (2007).

38. Debroy, D. M., Li, M., Huang, S. & Brown, K. R. Logical performance of 9 qubit compass codes in ion traps with crosstalk errors. *Quantum Sci. Technol.* **5**, 034002 (2020).

39. Li, M., Miller, D. & Brown, K. R. Direct measurement of Bacon-Shor code stabilizers. *Phys. Rev. A* **98**, 050301 (2018).

40. Li, M., Miller, D., Newman, M., Wu, Y. & Brown, K. R. 2D compass codes. *Phys. Rev. X* **9**, 021041 (2019).

41. Shor, P. Fault-tolerant quantum computation. *Proc. 37th Conf. on Foundations Comput. Sci.* (1996).

42. Terhal, B. M. Quantum error correction for quantum memories. *Rev. Mod. Phys.* **87**, 307 (2015).

43. Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *J. Math. Phys.* **43**, 4452–4505 (2002).

44. Reichardt, B. W. Quantum universality from magic states distillation applied to CSS codes. *Quantum Inf. Process.* **4**, 251–264 (2005).

45. Wang, P. *et al.* Single ion-qubit exceeding one hour coherence time. Preprint at https://arxiv.org/abs/2008.00251 (2020).

46. Eastin, B. & Knill, E. Restrictions on transversal encoded quantum gate sets. *Phys. Rev. Lett.* **102**, 110502 (2009).

47. Cetina, M. *et al.* Quantum gates on individually-addressed atomic qubits subject to noisy transverse motion. Preprint at https://arxiv.org/abs/2007.06768 (2020).

48. Kielpinski, D., Monroe, C. & Wineland, D. J. Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).

49. Home, J. P. *et al.* Complete methods set for scalable ion trap quantum information processing. *Science* **325**, 1227–1230 (2009).

50. Pino, J. *et al.* Demonstration of the QCCD trapped-ion quantum computer architecture. Preprint at https://arxiv.org/abs/2003.01293 (2020).

51. Brown, K. R., Harrow, A. W. & Chuang, I. L. Arbitrarily accurate composite pulse sequences. *Phys. Rev. A* **70**, 052318 (2004).

52. Mølmer, K. & Sørensen, A. Multiparticle entanglement of hot trapped ions. *Phys. Rev. Lett.* **82**, 1835 (1999).

53. Maslov, D. Basic circuit compilation techniques for an ion-trap quantum machine. *New J. Phys.* **19**, 023035 (2017).

54. Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).

55. Barends, R. *et al.* Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014).

56. Sackett, C. A. *et al.* Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000). Number: 6775 Publisher: Nature Publishing Group.

57. Inlek, I., Vittorini, G., Hucul, D., Crocker, C. & Monroe, C. Quantum gates with phase stability over space and time. *Phys. Rev. A* **90**, 042316 (2014).

58. Gullion, T., Baker, D. B. & Conradi, M. S. New, compensated carr-purcell sequences. *J. Magn. Reson. (1969)* **89**, 479–484 (1990).