nature portfolio

https://doi.org/10.1038/s41586-021-03928-y

Supplementary information

Fault-tolerant control of an error-corrected qubit

In the format provided by the authors and unedited

Supplementary Information: Fault-Tolerant Control of an Error-Corrected Qubit

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], and Christopher Monroe[1,2,3,6]

[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
[6]IonQ, Inc, 4505 Campus Drive, College Park, MD, 20740
[*]Present Address: Google Research, Venice, CA 90291
[†]Email: laird.egan@gmail.com, Present Address: IonQ, Inc, College Park, MD 20740

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[1] π-rotation to prepare $|1\rangle$. To measure the qubit state, 369 nm light that is resonant with the $\{^2\mathrm{S}_{1/2}, \mathrm{F}=1\} \longleftrightarrow \{^2\mathrm{P}_{1/2}, \mathrm{F}=0\}$ transition is directed 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.

Single qubit gate benchmarking

The reported single qubit gate fidelity was measured using single qubit randomized benchmarking[2], 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 $3.4(8) \times 10^{-4}$ on the 15-ion chain, corresponding to an error of $1.8(3) \times 10^{-4}$ per native Pauli gate[3]. The offset in the fit is consistent with SPAM errors.

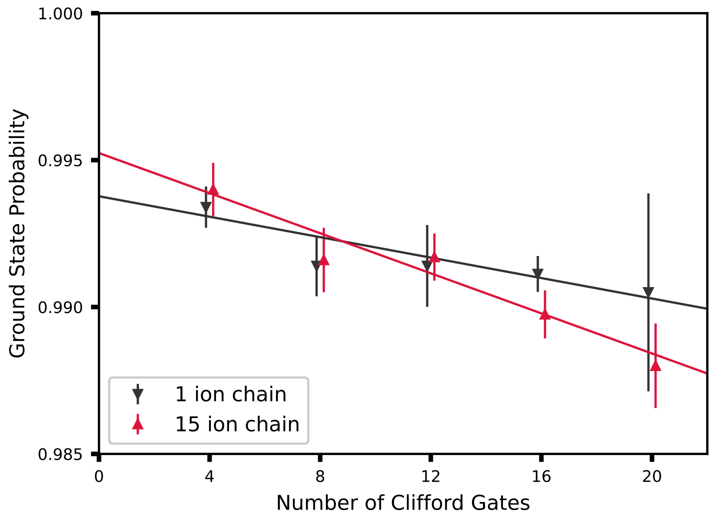

Figure S1. **Randomized benchmarking of single-qubit gates.** The probability to measure a single ion in the ground state after a variable number of Clifford gates in a randomized benchmarking sequence. The slope of the line indicates the per-Clifford fidelity, while the y-intercept indicates the SPAM 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.

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. We take the echoed sequence to be the "best-case" scenario and the non-echoed sequence to be the "worst-case" scenario. Within a circuit, we expect the true fidelity of a single XX gate to fall between these two extremes. The fidelity is determined by the parity fringe method[4]. 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. Thus, the estimated fidelity of a single XX gate is bounded within $98.5-99.3\%$. In other gates, a decrease in fidelity relative to this estimate is primarily due to effects of crosstalk.

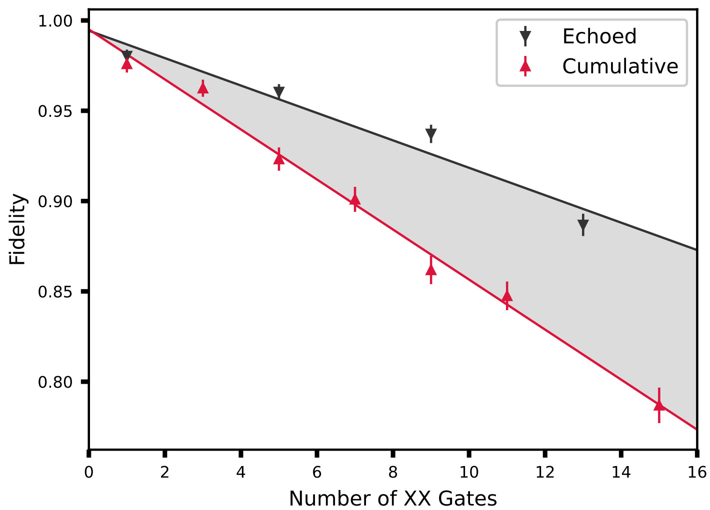

Figure S2. **Two-qubit gate performance.** The cumulative sequence (red) $XX(\pi/4+\varepsilon)XX(\pi/4+\varepsilon)...$ results in a total error of $N\varepsilon$, and the echoed sequence (black) $XX(\pi/4+\varepsilon)XX(-\pi/4-\varepsilon)...$ cancels this particular type of over/under-rotation error. 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.

Physical Error Modeling

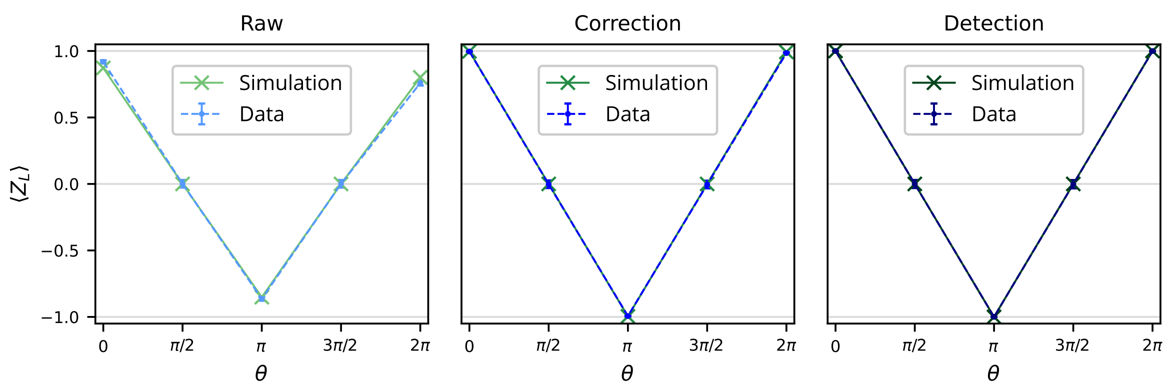

Figure S3. **Matching of simulation and experiment.** The simulated and experimentally measured data for repeated applications of the fault-tolerant logical rotation gate applied to a state initially prepared in $|0\rangle_L$. The data points labeled "data" are the same as the blue curves shown in Fig 3d of the main text.

To confirm our understanding of the experimental system, we find it useful to design physical error models that can replicate our experimental results in simulation. A simple coherent overrotation error model, combined with stochastic measurement and preparation errors closely matches our results, as shown in Fig. S3. The coherent overrotation error channel for a rotation gate G is modeled as

$$\varepsilon_G(\rho) = \exp(-i\theta G)\rho\exp(i\theta G), \tag{S1}$$

where θ is the angle of overrotation. This model is applied to both the Mølmer-Sørensen gate and the single qubit rotation gates, where the fidelities presented in the main text are directly translated into values for θ. These two errors are demonstrated in Fig. S4. The simulation then includes stochastic measurement errors and preparation errors following the SPAM error rates determined from our benchmarking experiments presented in the SPAM error benchmarking section above. Using this four parameter error model, we can capture much of the performance of our system.

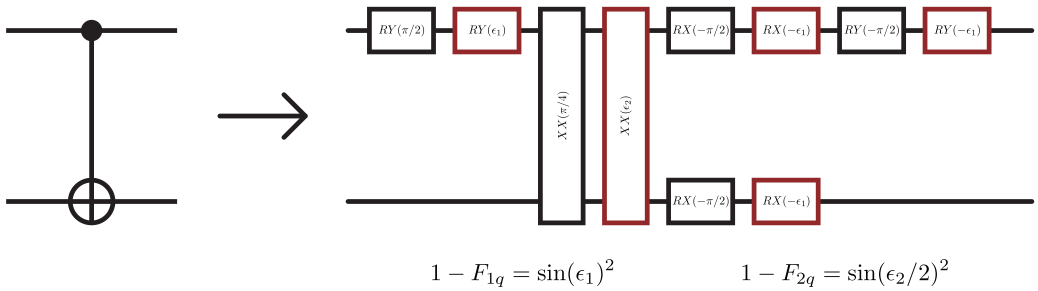

Figure S4. **Gate error model for CNOT.** One possible choice of native gate decomposition (black) and gate errors (red) for a CNOT gate.

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. S5, 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 ($\eta \approx 0.08$) converts carrier Rabi frequency to sideband frequency and this factor is normalized by the gate duration ($\tau_{gate} = 225\mu$ s). Using Fig. S5 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.

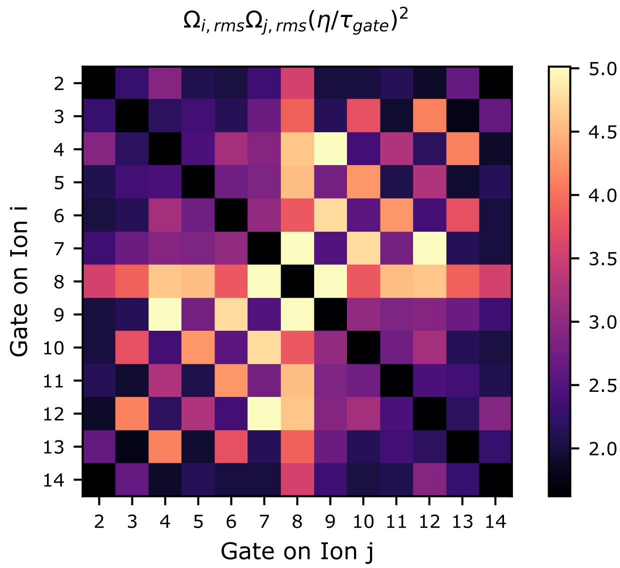

Figure S5. **Power requirements for XX gates in our system.** For an XX gate on ions i, j, the (RMS) Rabi frequency (Ω_{rms}) of the AM waveform for *each* red/blue sideband when brought into resonance with the carrier transition is normalized by the Lamb-Dicke factor ($\eta \approx 0.08$) and the gate 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.

Magic State Fidelity

To calculate the fidelity of our magic state preparation circuit for the state $|H_x\rangle_L$, we can compute the fidelity between a mixed state ρ, which represents the experimentally prepared state, and the ideal pure state $|\psi\rangle$ as

$$\begin{aligned} F &= \langle\psi|\rho|\psi\rangle \\ &= \mathrm{Tr}[\rho|\psi\rangle\langle\psi|] \\ &= \frac{1}{2}\left(1+\langle X\rangle_\rho\langle X\rangle_\psi+\langle Y\rangle_\rho\langle Y\rangle_\psi+\langle Z\rangle_\rho\langle Z\rangle_\psi\right) \\ &= \frac{1}{2}\left(1+\langle X\rangle_\rho\frac{1}{\sqrt{2}}+\langle Z\rangle_\rho\frac{1}{\sqrt{2}}\right). \end{aligned} \tag{S2}$$

The expectation values $\langle Z\rangle_\rho$ and $\langle X\rangle_\rho$ can be extracted by measuring logical Z operator before and after a logical $Y_L(\pi/2)$ operation. This analysis leads to the following fidelities as shown in Table S2.

Processing Technique	Fidelity
Raw	0.85(1)
Correction	0.97(1)
Detection	0.98(1)

Table S2. Fidelities for the $|H_x\rangle_L$ magic state preparation circuit under different processing techniques.

However the same procedure cannot be applied to the $|H_y\rangle_L$ state, as the [[9,1,3]] Bacon-Shor code does not allow for fault-tolerant measurement in the logical Y basis. Using the constraint

$$\langle X\rangle^2+\langle Y\rangle^2+\langle Z\rangle^2 \leq 1$$

we can only numerically bound the fidelity of the $|H_y\rangle_L$ state to the range $0.75 \leq F \leq 0.99$. However, we argue that the fidelities for preparing $|H_x\rangle_L$ and $|H_y\rangle_L$ should be very similar, as the preparation circuit only differs in the phase of a single qubit gate, a quantity which we control to $\approx 400\mu$rad limited by the AWG bit depth. This argument is further strengthened by the results of our single qubit benchmarking. Thus the $|H_y\rangle_L$ state fidelity should be very similar to values shown in S2.

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

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

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

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

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

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

In an experiment there will also be imperfections in the states due to errors beyond T_2^* dephasing, which we model to 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. S6.

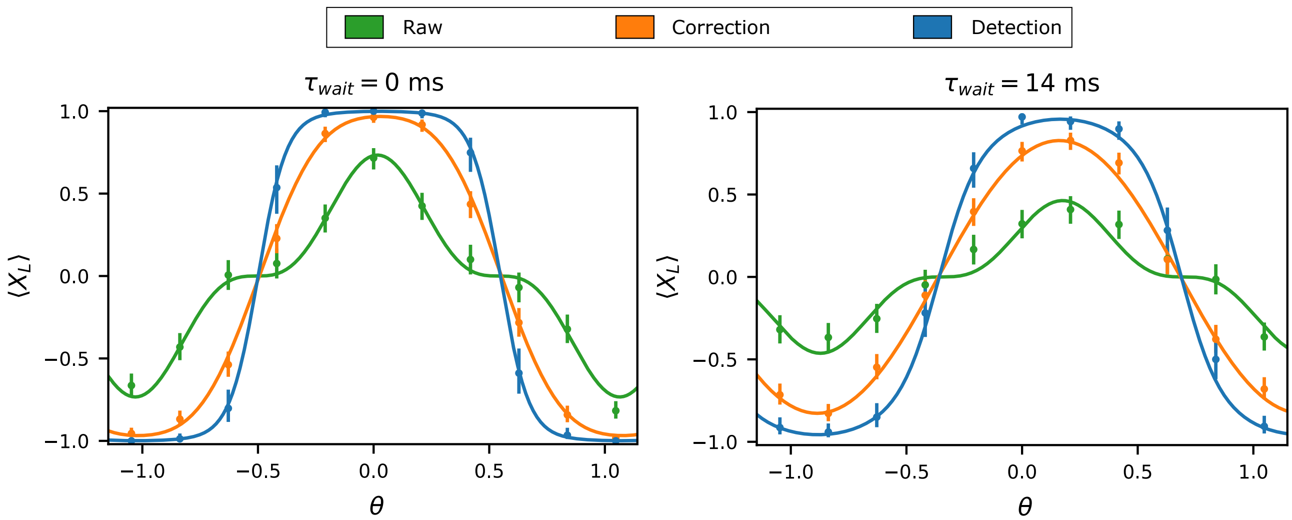

Figure S6. **Examples of T_2^* fits.** These plots exemplify logical Ramsey fringe fitting at two different wait times, 0ms (left) and 14ms (right). At short wait times (left), the data shows characteristics of error correction, such as the flattened top for error-detection. At longer wait times (right), 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

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. S7. 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[5].

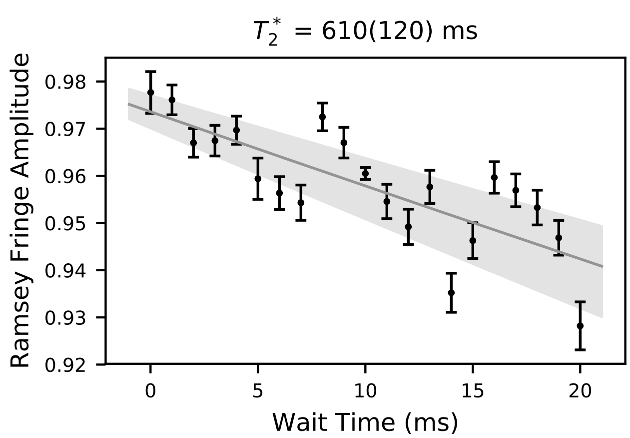

Figure S7. **Raman T_2^* for a physical qubit in a 15-ion chain.** Ramsey fringes with variable wait times are fit to a sinusoud to extract the contrast. For each data point, the error bars represent the 95% confidence interval from a maximal likelihood estimation fit of a sinuosoid. The contrast is 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.

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[6], 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. S8. 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 $^{171}\text{Yb}^+$[7].

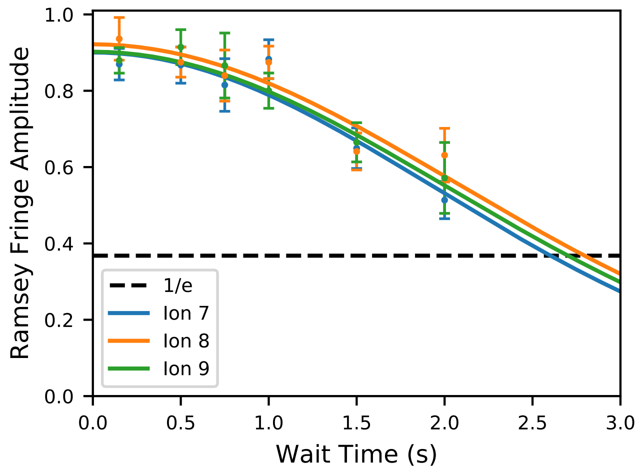

Figure S8. **Microwave T_2 for a physical qubit in a 15-ion chain.** An $(XY)^N$ dynamical decoupling sequence is used to suppress magnetic field noise. Ramsey fringes with variable wait times are fit to a sinusoud to extract the contrast. For each data point, the error bars represent the 1σ confidence 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.

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. S7, 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. S9, 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.

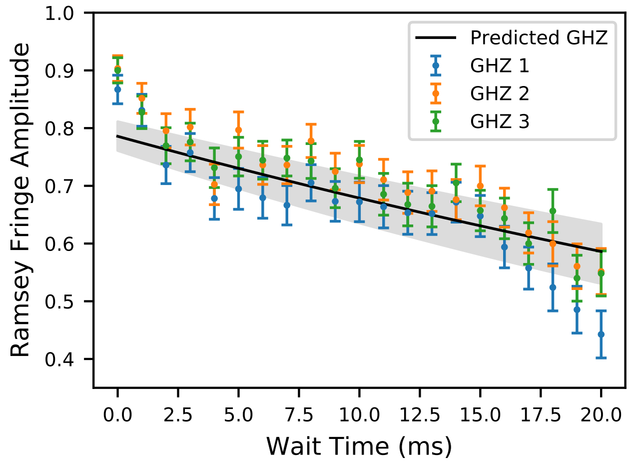

Figure S9. T_2^* **for GHZ states.** Ramsey fringes for each independent GHZ state are analyzed from the logical T_2^* experiment in Fig. 2c of the main text. For each data point, the error bars represent the 95% confidence interval from a maximal likelihood estimation fit of a sinuosoid. For the theoretical prediction, we apply the fitted dephasing noise in Fig. S7 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. S7, propagated through the numerical simulation.

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. S3 and S4, we report the numerical values obtained from fitting the logical gate operations as displayed in Fig. 4d of the main text. In addition to the FT transversal gate and the nFT continuous gate, we also provide the fit values for a similar sampling of states in the logical XZ-plane generated from the nFT encoding circuit (Fig. 2a of 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	0.915(6)	0.995(1)	$1.000(^{0}_{3})$
nFT Gate	0.87(1)	0.87(1)	$1.00(^{0}_{1})$
nFT Encoding	0.78(1)	0.982(5)	0.997(2)

Table S3. Numerical values for the fit parameter A.

Gate Error (Γ)	Raw	Error Correction	Error Detection
FT Gate	0.045(3)	0.0027(7)	0.0002(1)
nFT Gate	0.020(5)	0.000(4)	0.011(2)
nFT Encoding	0.001(6)	0.000(2)	0.001(1)

Table S4. Numerical values for the fit parameter Γ.

Stabilizers on different input states

In the main text (Fig 3a), we describe an experiment to measure X-type errors on the logical qubit state propagated from a Z-type errors on an ancilla during a FT X stabilizer measurement circuit. In that experiment, there was no detectable change in the error rate. However, because the input and measured state was $|0\rangle_L$, the measurement is only sensitive to X-type errors. To check if the stabilizer is introducing Z-type errors into the logical qubit, we perform the same experiment (with no artificial error added) on the $|+\rangle_L$ input state. $|+\rangle_L$ is measured in the $\langle X\rangle_L$ basis by a transversal $Y_L(-\pi/2)$ gate that maps $\langle X\rangle_L \rightarrow \langle Z\rangle_L$. We also check a Z stabilizers on both input states. The results are shown in table S5. We note that from the logical T_2^* experiment, we expect the error on the $|+\rangle_L$ state to increase by 1.8% over the $\approx$ 1.5ms required to measure the stabilizer, which is not included in the baseline encoding error below.

Input State	Baseline Encoding (% Error)	FT Z Stabilizer (% Error)	FT X Stabilizer (% Error)
$\|0\rangle_L$	0.23(13)	0.42(10)	0.25(14)
$\|+\rangle_L$	0.44(10)	3.4(3)	2.1(2)

Table S5. Logical error rates for both Z and X FT stabilizers on different logical input states, after error-correction. Uncertainties are 1σ from the binomial distribution.

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. S10, we present a full set of errors that produce all of the possible ancilla qubit output bit strings.

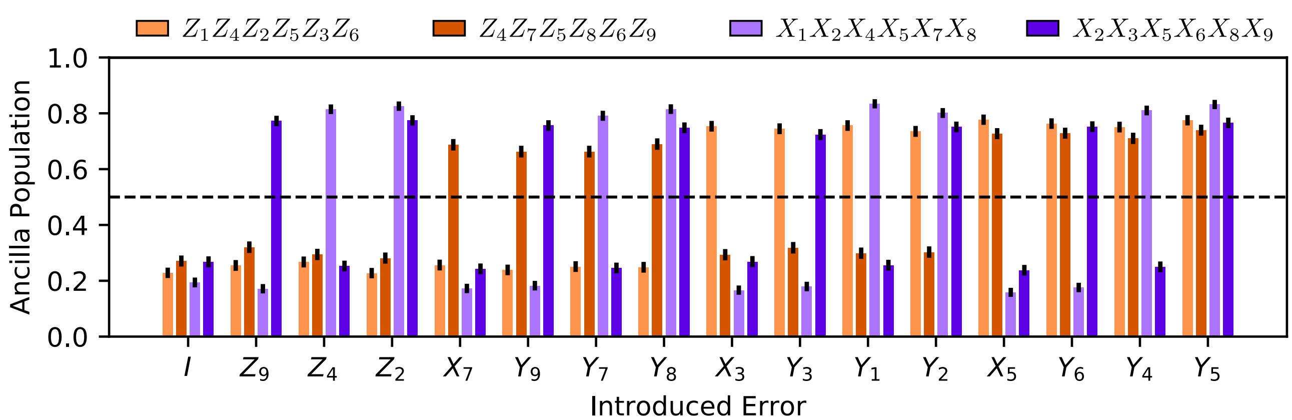

Figure S10. **Full stabilizer measurement results.** A complete set of errors are introduced to create all possible output bit strings of the ancilla qubits. Error bars are the 95% binomial proportion confidence interval.

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:

$$\begin{aligned}
\varepsilon_{S_1} &= \varepsilon_{enc} + 2\varepsilon_X + \varepsilon_Z \\
\varepsilon_{S_2} &= \varepsilon_{enc} + 2\varepsilon_X + 2\varepsilon_Z \\
\varepsilon_{S_3} &= \varepsilon_{enc} + \varepsilon_{T_2^*} + \varepsilon_X \\
\varepsilon_{S_4} &= \varepsilon_{enc} + \varepsilon_{T_2^*} + 2\varepsilon_X
\end{aligned} \tag{S3}$$

The stabilizer eigenvalues obtained from the ancilla qubits are useful for error-correction if they are well correlated with the true stabilizer eigenvalues obtained from the physical qubits that encode the logical qubit. We quantify this correlation by calculating the covariance between these two measurement methods. If no error has occurred during or after mapping the stabilizer to the ancilla, we should expect the ancilla to always agree with the data qubits and have unity covariance. Defining the stabilizer eigenvalues obtained from the ancilla qubits as $S_{1,a}$ and $S_{2,a}$ (the data shown in Figure S10) and the stabilizer eigenvalues obtained from the data qubits at the end of the circuit as $S_{1,d}$ and $S_{2,d}$ (the process we use for offline error correction), we observe a covariance of $\mathrm{cov}(S_{1,a}, S_{1,d}) = 0.80$ and $\mathrm{cov}(S_{2,a}, S_{2,d}) = 0.82$.

Native Gate Circuit Library

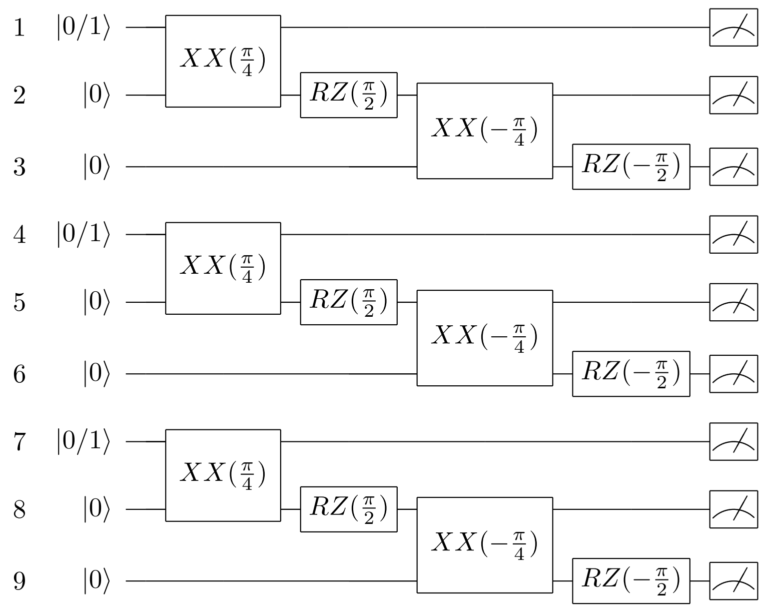

Figure S11. **Native gate compilation for encoding $|0/1\rangle_L$.** To create the $|0/1\rangle_L$ state, qubits 1, 4, and 7 should be initialized in the $|0/1\rangle$ state, respectively. Corresponds to the blue Clifford circuit in Fig. 2a and the blue data in Fig. 2b of the main text/

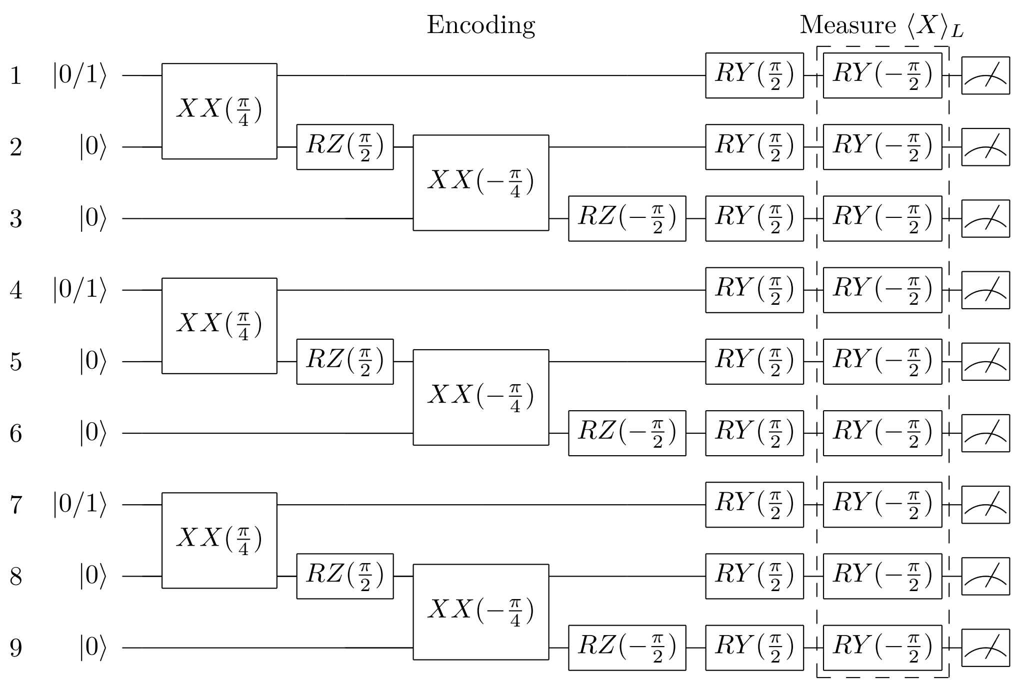

Figure S12. **Native gate compilation for encoding $|+/-\rangle_L$.** To create the $|+/-\rangle_L$ state, qubits 1, 4, and 7 should be initialized in the $|0/1\rangle$ state, respectively. We note that we do not compile out the successive $RY(\pm\frac{\pi}{2})$ gates on each qubit, as this would create a circuit identical to S11 - this explains the additional error on $|+/-\rangle_L$ states relative to $|0/1\rangle_L$. Corresponds to the blue circuit in Fig. 2a and the blue data in Fig. 2b of the main text.

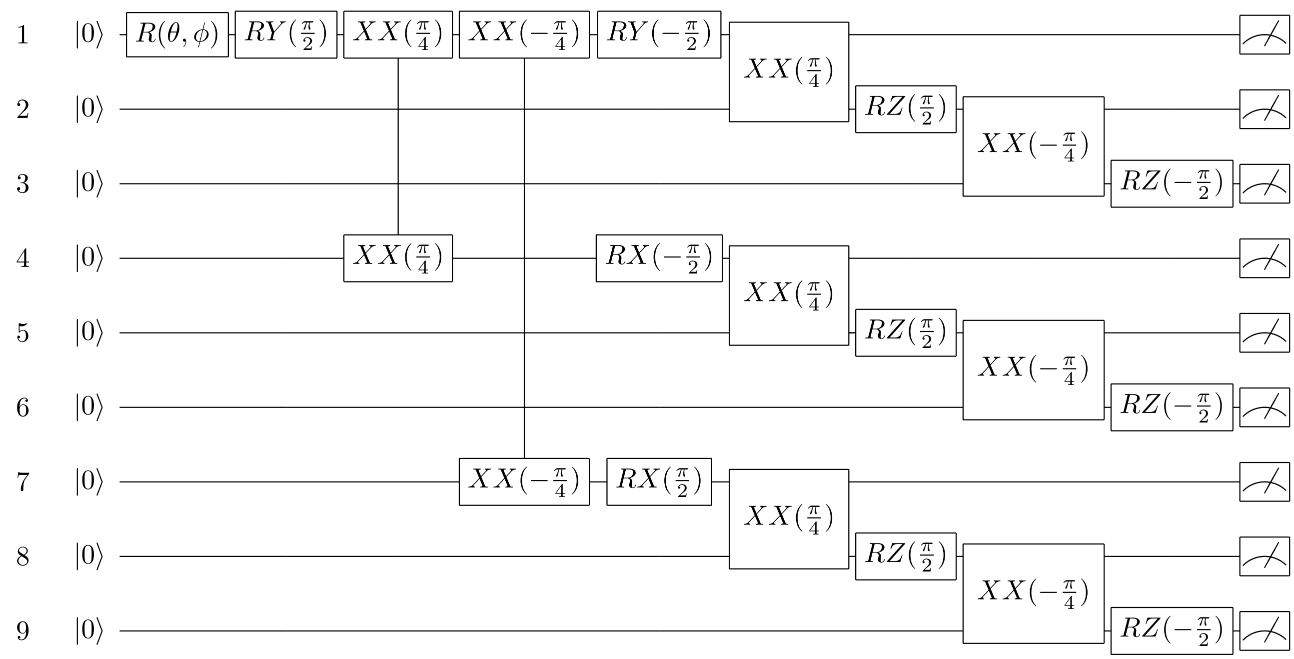

Figure S13. **Native gate compilation for nFT encoding of arbitrary logical states.** This circuit can create an arbitrary $|\psi\rangle_L$ state controlled by the initial $R(\theta,\phi)$ gate on qubit 1, with $|\psi\rangle_1 = R(\theta,\phi)|0\rangle_1$. Corresponds to the red circuit in Fig. 2a and the red data in Fig. 2b of the main text. This circuit is also used for initial encoding of the magic states in Fig. c-d of the main text.

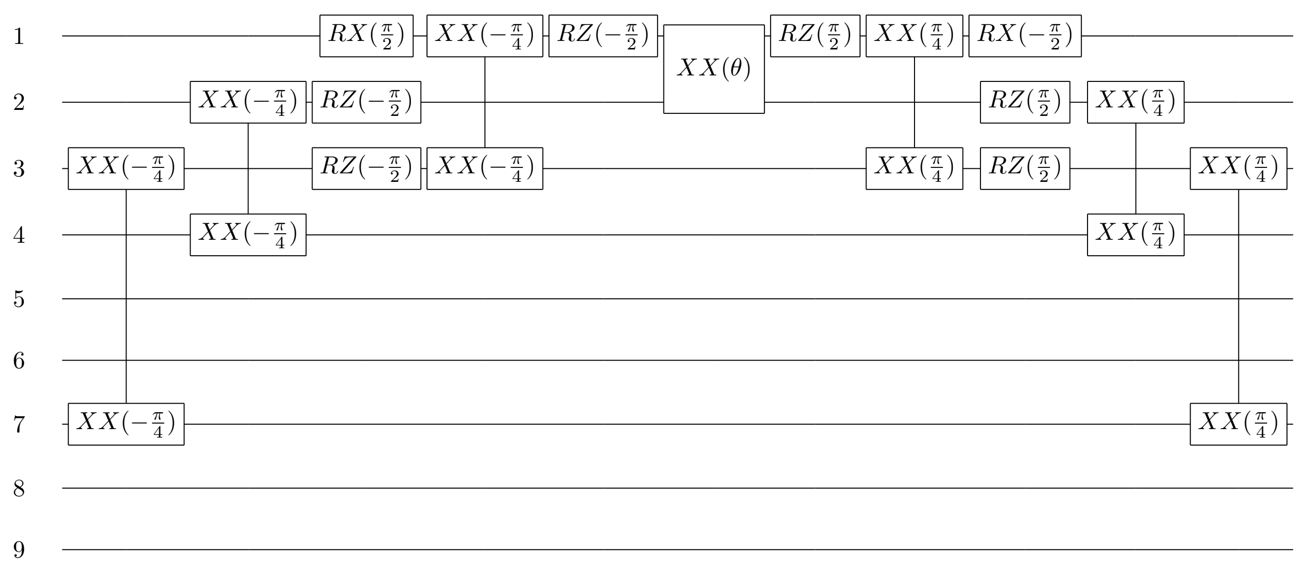

Figure S14. **Native gate compilation for nFT continuous rotation.** This circuit implements a $Y_L(\theta) = Y_1 Z_2 Z_3 X_4 X_7(\theta)$ rotation controlled by an $XX(\theta)$ gate on qubits 1 and 2. Corresponds to the red circuit in Fig. 3b and the red data in Fig. 3d-e of the main text.

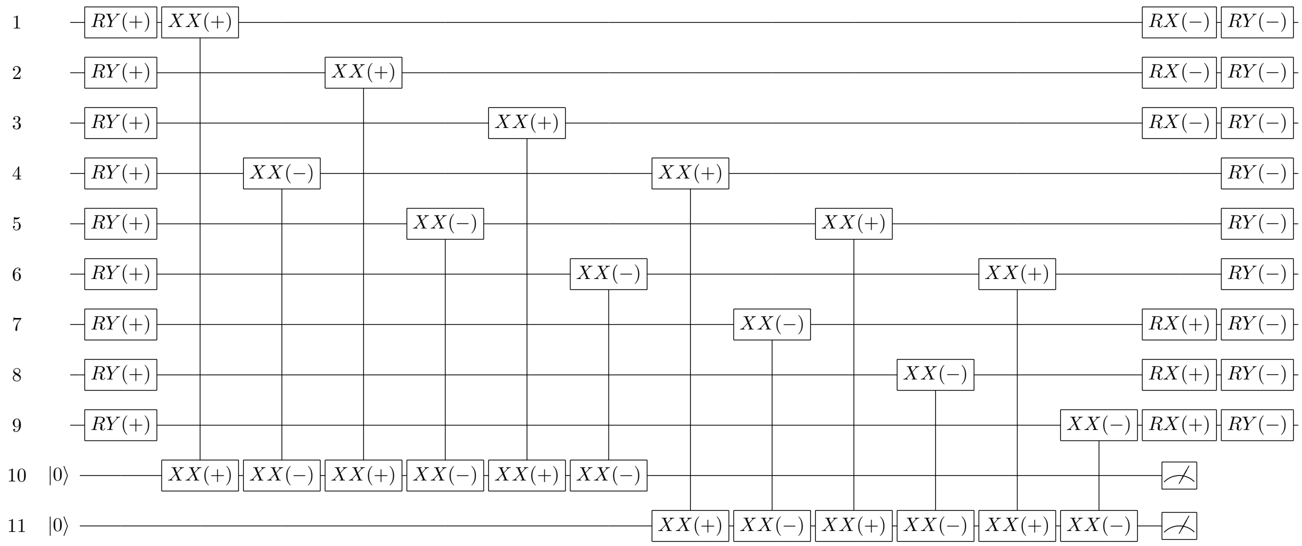

Figure S15. **Native gate compilation for Z stabilizer measurement.** Here, the angle $(+/-)$ is shorthand for $(\pm\frac{\pi}{2})$ or $(\pm\frac{\pi}{4})$ for single- or two-qubit gates, respectively. When combined with encoding (Fig. S11) and the X stabilizer (Fig. S16), the combined circuits measure the full syndrome (Extended Data Fig. 1a).

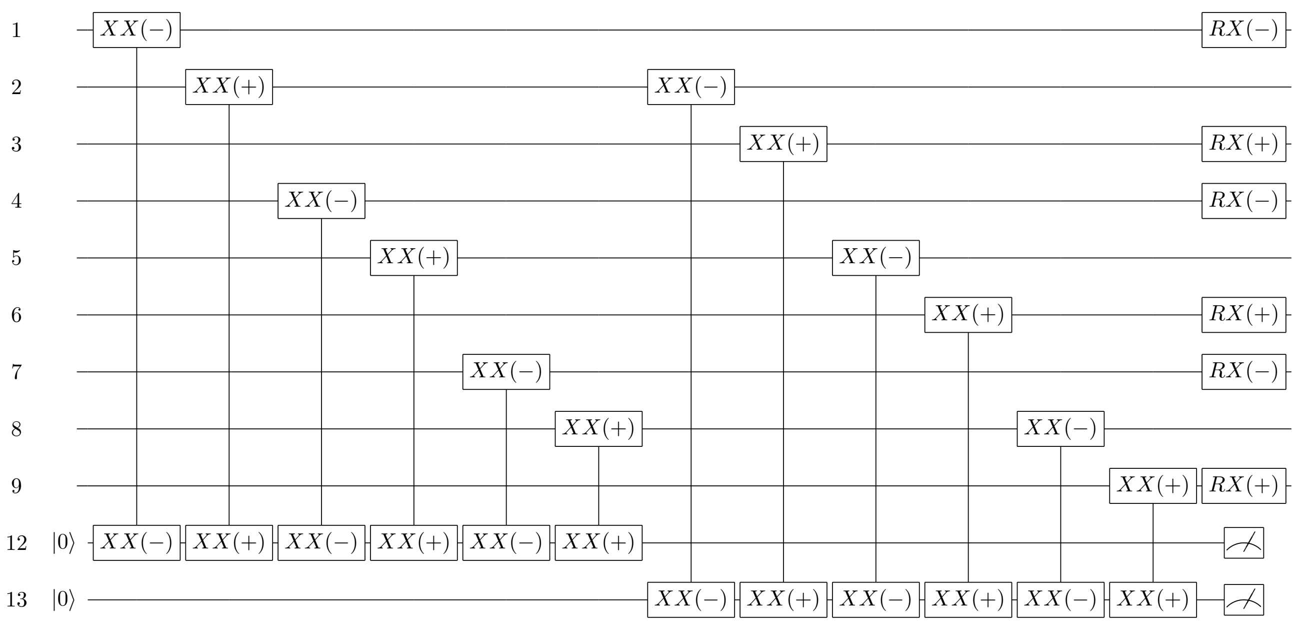

Figure S16. **Native gate compilation for FT X stabilizer measurement.** Here, the angle $(+/-)$ is shorthand for $(\pm\frac{\pi}{2})$ or $(\pm\frac{\pi}{4})$ for single- or two-qubit gates, respectively. Corresponds to the circuit in Extended Data Fig. 1c and for data presented in Fig. 4a of the main text.

References

1. Brown, K. R., Harrow, A. W. & Chuang, I. L. Arbitrarily accurate composite pulse sequences. *Phys. Rev. A* **70**, 052318 (2004).
2. Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
3. Barends, R. *et al.* Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014).

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

5. 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).

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

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