Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/12/eaay3050/DC1

Supplementary Materials for

Protecting quantum entanglement from leakage and qubit errors via repetitive parity measurements

C. C. Bultink, T. E. O'Brien, R. Vollmer, N. Muthusubramanian, M. W. Beekman, M. A. Rol, X. Fu, B. Tarasinski, V. Ostroukh, B. Varbanov, A. Bruno, L. DiCarlo*

*Corresponding author. Email: l.dicarlo@tudelft.nl

Published 20 March 2020, *Sci. Adv.* **6**, eaay3050 (2020)
DOI: 10.1126/sciadv.aay3050

This PDF file includes:

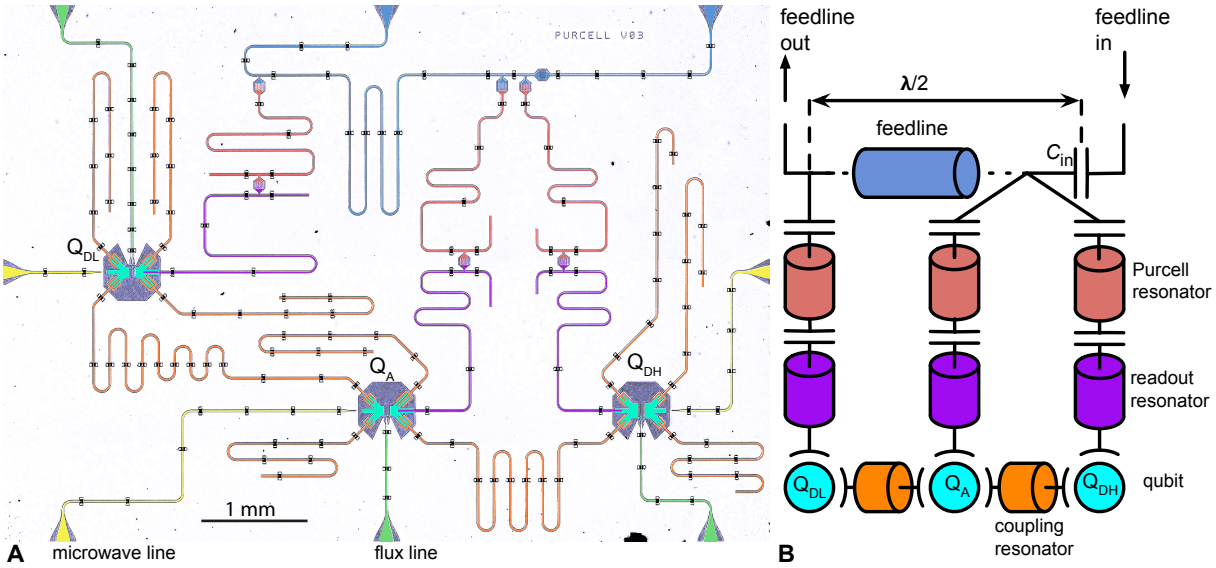

Fig. S1. False-colored photograph (**A**) and simplified circuit diagram (**B**) of the quantum processor with corresponding colors.

Supplementary Materials and Methods

1.1 Setup

A full wiring diagram of the setup is provided in (Fig. S2). All operations are controlled by a fully digital device, the central controller (CC7), which takes as input a binary in an executable quantum instruction set architecture [eQASM (*46*)], and outputs digital codeword triggers based on the execution result of these instructions. These digital codeword triggers are issued every 20 ns to arbitrary waveform generators (AWGs) for single-qubit gates and two-qubit gates, a vector switch matrix (VSM) for single-qubit gate routing and a readout module (AWG and acquisition) for frequency-multiplexed readout. Single-qubit gate generation, readout pulse generation and readout signal integration are performed by single-sideband mixing. The measurement signal is amplified with a JTWPA (*47*) at the front end of the amplification chain. Following Ref. (*48*), we extract an overall measurement efficiency $\eta = 48 \pm 1.0\%$ by comparing the integrated signal-to-noise ratio of single-shot readout to the integrated measurement-induced dephasing.

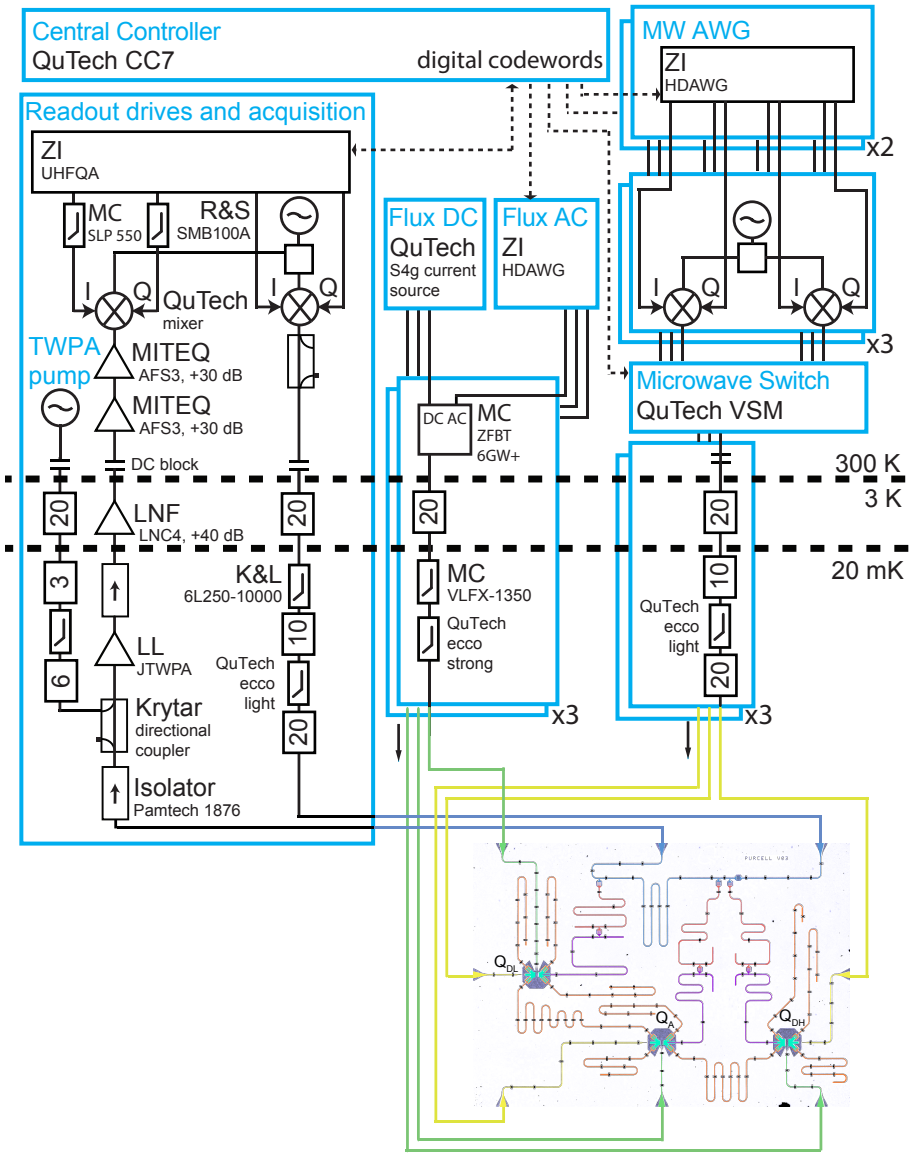

Fig. S2. Complete wiring diagram of electronic components inside and outside the ^{3}He/^{4}He dilution refrigerator (Leiden Cryogenics CF-CS81).

1.2 Cross-measurement-induced dephasing of data qubits

During ancilla measurement, data-qubit coherence is susceptible to intrinsic decoherence, phase shifts via residual ZZ interactions and cross-measurement-induced dephasing (*8, 33*). For the single-data-qubit subspace we investigate the different contributions experimentally and assess the benefit of an echo pulse on the data qubits halfway through the ancilla measurement. We study this by including the ancilla measurement (with amplitude ε) in a Ramsey-type sequence (Fig. S3A). By varying the azimuthal phase of the second $\pi/2$ pulse, we obtain Ramsey fringes from which we extract the coherence $|\rho_{01}|$ and phase $arg\,(\rho_{01})$. Several features of these curves explain the need for the echo pulse on the data qubits. Firstly, at $\varepsilon = 0$, the echo pulse improves data-qubit coherence (for both ancilla states) by reducing the effect of low-frequency noise (Fig. S3, B and C). This is confirmed by individual Ramsey and echo experiments. Secondly, the echo pulse almost perfectly cancels ancilla-state dependent phase shifts due to residual ZZ interactions (Fig. S3, D and E). When gradually turning on the ancilla measurement towards the nominal value $\varepsilon = 1$, we furthermore observe that: thirdly, the echo pulse almost perfectly cancels the measurement-induced Stark shift (Fig. S3, D and E). When increasing the measurement amplitude beyond the operation amplitude (indicated by the vertical dashed lines), we see rapid non-Gaussian decay of data-qubit coherence. We attribute this to measurement-induced relaxation of the ancilla: via the ZZ interaction, this can lead to probabilistic phase shifts on the data qubit. This effect is stronger for Q_{DL} than for Q_{DH} due to its higher residual interaction with Q_{A} (Table S1).

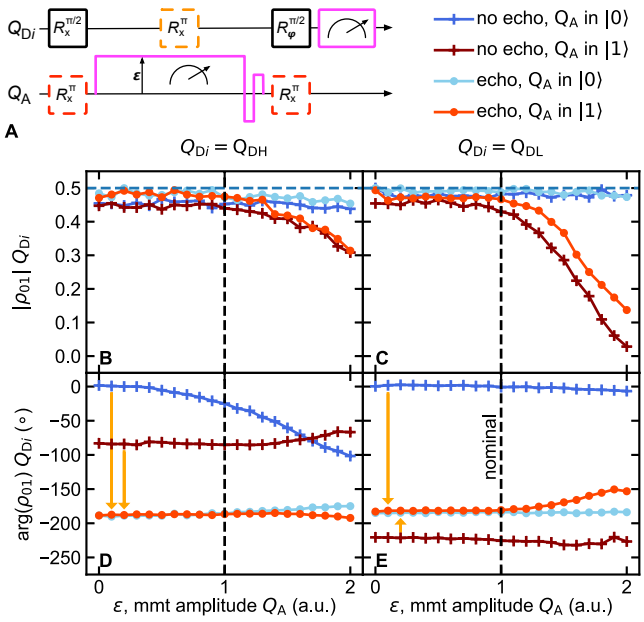

Fig. S3. Study of data-qubit coherence and phase accrual during ancilla measurement. (**A**) Quantum circuit to extract data-qubit coherence and phase with or without echo pulse (orange) and with or without excitation in the ancilla. (**B** and **C**) Data-qubit coherence as a function of ancilla measurement amplitude. (**D** and **E**) Data-qubit phase as a function of ancilla measurement amplitude.

Table S1. Measured parameters of the three-transmon device. † Q_{DH} is operated 30 MHz below its maximum frequency to avoid spurious interaction with a spurious two-level system. †† The Purcell mode and readout resonator mode of Q_{DH} have near-perfect hybridization (with qubit at $\omega_{\mathrm{op}}/2\pi$), making them indistinguishable. ††† Single-shot readout on the data qubits was not optimized. †††† Single-qubit gates are characterized using Clifford randomized benchmarking (*49*) ††††† Two-qubit gates are characterized using interleaved RB (*2, 49*) with a leakage-extraction modification (*5*).

Gate and Coherence Parameters	Q_{DL}	Q_{A}	Q_{DH}
operating qubit frequency, $\omega_{\mathrm{op}}/2\pi$ (GHz)	5.02	5.79	6.88†
max. qubit frequency, $\omega_{\mathrm{max}}/2\pi$ (GHz)	5.02	5.79	6.91
anharmonicity, $\alpha/2\pi$ (MHz)	-306	-308	-331
coherence time (at $\omega_{\mathrm{op}}/2\pi$), T_2^{echo} (μs)	29.6 ± 2.7	21.7 ± 1.4	14.7 ± 0.9
relaxation time (at $\omega_{\mathrm{op}}/2\pi$) T_1 (μs)	25.3 ± 1.2	17.0 ± 0.6	25.6 ± 1.2
Ramsey dephasing time (at $\omega_{\mathrm{op}}/2\pi$), T_2^* (μs)	24.5 ± 2.0	14.6 ± 1.2	5.9 ± 0.7
average error per single qubit gate††††, e_{SQ} (%)	0.08 ± 0.02	0.14 ± 0.016	0.21 ± 0.06
resonance exchange coupling, $J_1/2\pi$ (MHz)	17.2	14.3	
bus resonator frequency, $\sim \omega_{\mathrm{bus}}/2\pi$ (GHz)	8.5	8.5	
error per CZ†††††, e_{CZ} (%)	1.4 ± 0.6	0.9 ± 0.16	
leakage per CZ†††††, L_1 (%)	0.27 ± 0.12	0.15 ± 0.07	
ZZ coupling (at $\omega_{\mathrm{op}}/2\pi$), $\zeta_{\mathrm{ZZ}}/2\pi$ (MHz)	0.95	0.33	
Measurement Parameters	Q_{DL}	Q_{A}	Q_{DH}
readout pulse frequency, $\omega_{\mathrm{ro}}/2\pi$ (GHz)	7.225	7.420	7.838
readout resonator frequency, $\omega_{\mathrm{ro}}/2\pi$ (GHz)	7.275	7.385	7.867
Purcell resonator frequency, $\omega_{\mathrm{ro}}/2\pi$ (GHz)	7.260	7.405	7.872
qubit-RR coupling strength, $g_{01,\mathrm{RR}}/2\pi$ (MHz)	202	188	135
PF-RR coupling strength, $J_{\mathrm{RR,PF}}/2\pi$ (MHz)	48	30	38
dispersive shift qubit-RR, χ_{RR}/π (MHz)	-2.5	-5.3	-2.8††
dispersive shift qubit-PF, χ_{PF}/π (MHz)	-1.5	-4.7	-2.8††
critical photon number, n_{crit}	2.3	2.7	2.4
intra-resonator photon number RR, n_{RR}		1.2	
quantum efficiency, η (%)		48 ± 1.0	
Average assignment error, e_{a} (%)	9.0†††	1.0 ± 0.1	16†††
Measurement integration time, τ_{int} (ns)	600	600	600

Supplementary Text

2.1 Performance of the simple hidden Markov model

In this section we detail the performance of the simple HMMs, as described in Fig. 3A and Sec. 4.4 of the main text. In Fig. S10, A and B, we plot a histogram of the computational likelihoods $L_{\mathrm{comp,Q}}$ of 10^5 simulated and actual ZZ experiments as calculated with the simple HMMs $H_{\mathrm{ZZ}} - D^{(\mathrm{simple})}$ and $H_{\mathrm{ZZ}} - A^{(\mathrm{simple})}$. This can be compared with Fig. 3, B and C, of the main text. We plot similar histograms for the interleaved ZZ—XX experiment in Fig. S10, H and I. We see reasonable agreement, but noticeably worse agreement than that in the detailed model. This is underscored by the Akaike information criterion (Eq.1 of the main text), which is significantly reduced compared to the more detailed HMMs

$$A(H_{\mathrm{ZZ}} - D) - A(H_{\mathrm{ZZ}} - D^{(\mathrm{simple})}) = 4.5 \times 10^5 \tag{S1}$$

$$A(H_{\mathrm{ZZ}} - A) - A(H_{\mathrm{ZZ}} - A^{(\mathrm{simple})}) = 5.9 \times 10^6 \tag{S2}$$

$$A(H_{\mathrm{ZZ,XX}} - D) - A(H_{\mathrm{ZZ,XX}} - D^{(\mathrm{simple})}) = 1.5 \times 10^5 \tag{S3}$$

$$A(H_{\mathrm{ZZ,XX}} - A) - A(H_{\mathrm{ZZ,XX}} - A^{(\mathrm{simple})}) = 1.6 \times 10^6 \tag{S4}$$

Indeed, in all cases the Akaike information criterion for the simple HMM is lower than that for the detailed HMM without leakage. This makes complete sense, as even though the simple HMMs might capture leakage fairly well, the additional effects captured in the detailed HMMs are far more dominant in the measurement signals than that of leakage. As such, the internal metrics, such as the ROC curves (Fig. S11) for the simplified model are significantly less trustworthy than those of the detailed model. This exemplifies the need for external HMM verification, as achieved in the main text by testing the HMM in a leakage mitigation scheme. We now repeat this verification procedure for the simple model. We see that in the ZZ experiment the performance is significantly degraded; although the flat line in the $\langle Z \otimes Z \rangle$ curve is restored

after about 8 parity checks, it requires rejecting 47% of the data, and is restored to a point $\sim 8\%$ below the performance of the detailed HMM. By contrast, the simple HMM performs almost identically to the complex HMM in the interleaved ZZ—XX experiment, achieving Bell-state fidelities within 2% whilst retaining the same amount of data. As the signal from a large-scale QEC code is more similar to the latter experiment than the former (See Sec. 2.2), this strongly suggests that the detailed modeling used in this text will not be needed in such experiments.

2.2 Hidden Markov models for large-scale QEC

The hidden Markov models used in this text provide an exciting prospect for the indirect detection of leakage on both data qubits and ancillas in a QEC code. This is essential for accurate decoding of stabilizer measurements made during QEC. Furthermore, this idea can be combined with proposals for leakage reduction (*14–17*) to target such efforts, reducing unnecessary overhead. As leakage does not spread in superconducting qubits (to lowest order), and gives only local error signals (*16*), such a scheme would require a single HMM per (data and ancilla) qubit. Each individual HMM needs only to process the local error syndrome, and as demonstrated in this work, completely independent HMMs may be used for the detection of nearby data-qubit and ancilla leakage. This implies that the computational overhead of leakage detection via HMMs in a larger QEC code will grow only linearly with the system size. Previous leakage reduction units are designed to act as the identity on the computational subspace (up to additional noise), so we do not require perfect discrimination between leaked and computational states. However, optimizing this discrimination (and investigating threshold levels for the application of targeted leakage reduction) will boost the code performance. Also, near-perfect discrimination could allow for the direct resetting of leaked data qubits (*40*), which would completely destroy an error correcting code if not targeted.

On the other hand, for implementation on classical hardware within the sub-$1\ \mu s$ QEC cycle

time on superconducting qubits (*36*), one may wish to strip back some of the optimization used in this work. The minimal HMM that could be used in QEC for detection has only two states, leaked and unleaked (Fig. 3A), and 2^{n_A} outputs, where n_A is the number of neighboring ancilla on which a signature of leakage is detected. (For the surface code, $n_\mathrm{A} \leq 4$ in all situations.) Such a simple model cannot perfectly deal with correlated errors, such as ancilla errors (which give multiple error signals separated in time). However, this should only cause a slight reduction in the discrimination capability whenever such correlations remain local. If the loss in accuracy is acceptable, one may store only $\pi_0^{(\mathrm{post})}$, and update it following a measurement $M_\mathrm{A}[m]$ as

$$\pi_0^{(\mathrm{prior})}[m] = (A_{0,0} - A_{0,1})\pi_0^{(\mathrm{post})}[m-1] + A_{0,1} \tag{S5}$$

$$\pi_0^{(\mathrm{post})}[m] = \frac{\pi_0^{(\mathrm{prior})}[m] B_{M_\mathrm{A}[m],0}}{B_{M_\mathrm{A}[m],1} + \pi_0^{(\mathrm{prior})}(B_{M_\mathrm{A}[m],0} - B_{M_\mathrm{A}[m],1})} \tag{S6}$$

which is trivial compared to the overhead for most QEC decoders.

A key question about the use of HMMs for leakage detection in future QEC experiments is whether leakage in larger codes is reliably detectable. In previous theoretical work (*50*), data-qubit leakage in repetition codes has been sometimes hidden, a phenomenon known as 'leakage paralysis' or 'silent stabilizer' (*51*). This effect occurs when the relative phase φ accumulated between the $|20\rangle$ and $|21\rangle$ states during a CZ gate is a multiple of π. In the absence of additional error, an indirect measurement of the data qubit via an ancilla would return a result $\frac{\varphi}{\pi} \bmod 2$. (By comparison, if $\varphi = \pi/2$, the ancilla would return measurements of 0 or 1 at random.) This is then identical to the measurement of a data qubit in the $|\frac{\varphi}{\pi} \bmod 2\rangle$ state, and no discrimination between the two may be achieved. However, in an N-qubit parity check S, the ancilla continues to accumulate phase from the other qubits, reducing this to an $N-1$-qubit effective parity check S' (plus a well-defined, constant phase). Such a parity check may no longer commute with other effective parity checks R' that share the leaked qubit, even though we would require $[S, R] = 0$ in stabilizer QEC. This is demonstrated in our second experiment measuring both ZZ and XX

parity checks; though these commute when no data qubit is leaked, leakage reduces the checks to non-commuting Z and X measurements (of the unleaked data qubit). (In the ZZ experiment, the leakage paralysis was broken by the echo pulse on the data qubits, which flips the effective stabilizer of a leaked qubit at each round.) The repeated measurement of these non-commuting operators generates random results, similar to the case when $\varphi = \pi/2$. To the best of our knowledge, in all fully fault-tolerant stabilizer QEC codes, the removal of a single data qubit breaks the commutativity of at least two neighboring stabilizers. As such, data-qubit leakage will always be detectable in QEC experiments with superconducting circuits.

Beyond the proof-of-principle argument above, one might question whether the signal of leakage is improved or reduced when going from our prototype experiment to a larger QEC code, and when the underlying physical-qubit error rate is reduced. Fortunately, we can expect an improvement in the HMM discrimination capability in both situations. To see this, consider the example of a data qubit which is either leaked at round 1 with probability $p_{\rm leak}$ or never leaks. Let us further assume that in the absence of leakage, a number of neighboring ancillas $n_{\rm A}$ incur errors (where the parity check reports a flip) at a rate p, whereas in the presence of leakage these ancillas incur errors at a rate 0.5. (For example, in the bulk of the surface code, $n_{\rm A} = 4$.) The computational likelihood at round $m > 0$ after seeing e errors may be calculated as

$$L_{\rm comp,Q}[m] = \frac{(1-p_{\rm leak})p^e(1-p)^{mn_{\rm A}-e}}{(1-p_{\rm leak})p^e(1-p)^{mn_{\rm A}-e} + p_{\rm leak}(0.5)^{mn_{\rm A}}} \tag{S7}$$

If the data qubit was leaked, $e \sim mn_{\rm A}/2$, and the computational likelihood on average is approximately

$$L_{\rm comp}[m] \sim \frac{1-p_{\rm leak}}{p_{\rm leak}} \left(\frac{p^{n_{\rm A}/2}(1-p)^{n_{\rm A}/2}}{0.5^{n_{\rm A}}} \right)^m \tag{S8}$$

which is of the form

$$L_{\rm comp}[m] = Ae^{-\lambda m}, \quad A = \frac{1-p_{\rm leak}}{p_{\rm leak}}, \quad \lambda = \log\left(2^{n_{\rm A}} p^{-\frac{n_{\rm A}}{2}} (1-p)^{-\frac{n_{\rm A}}{2}}\right) \tag{S9}$$

We see that the signal of leakage ($L_{\rm comp}[m] \to 0$) switches on exponentially in time, with a rate proportional to $\log(p^{-n_{\rm A}/2})$. Any decrease in p (from better qubits) or increases in $n_{\rm A}$ (from additional ancillas surrounding the leaked qubit in a QEC code) will serve to increase, and not decrease this rate. The exponential decay constant is inversely proportional to the leakage rate (as this corresponds to an initial HMM skepticism towards unlikely leakage events). However, as the likelihood 'switch' is exponential, a decrease in $p_{\rm leak}$ by even an order of magnitude should only increase the time before definite detection by a single step or so. The above analysis is complicated in a real scenario, as single physical errors give correlated detection signals, and as leakage may occur at any time, and as leaked qubits may seep. Correlations in the detection signals will serve to renormalize the switching time λ (but not remove the generic feature of exponential onset). Seepage causes individual leakage events to be finite (with some average lifetime $T_{\rm seep}$); an individual leakage event of length $\ll \lambda^{-1}$ will not be detectable by the HMM. However, when the system returns to the computational subspace in such a short period of time, the leakage event may be treated as a 'regular' error, and does not need complicated leakage-detection hardware for fault tolerance. For example, a leakage event followed by immediate decay to $|1\rangle$ is indistinguishable from a direct transition to $|1\rangle$ for all practical purposes in QEC.

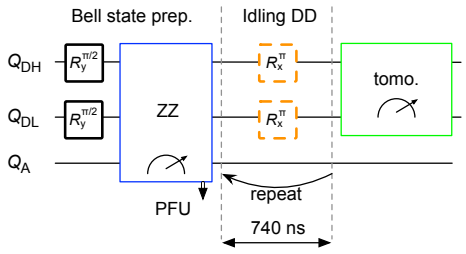

Fig. S4. Quantum circuit for Bell-state idling experiments under dynamical decoupling.

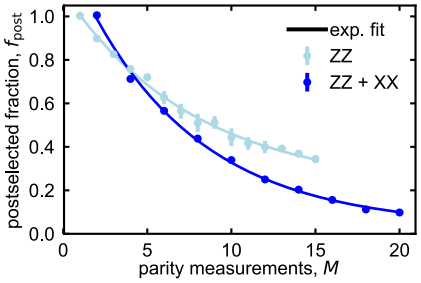

Fig. S5. Postselected fractions for the "no error" conditioning in Figs. 2 and 4.

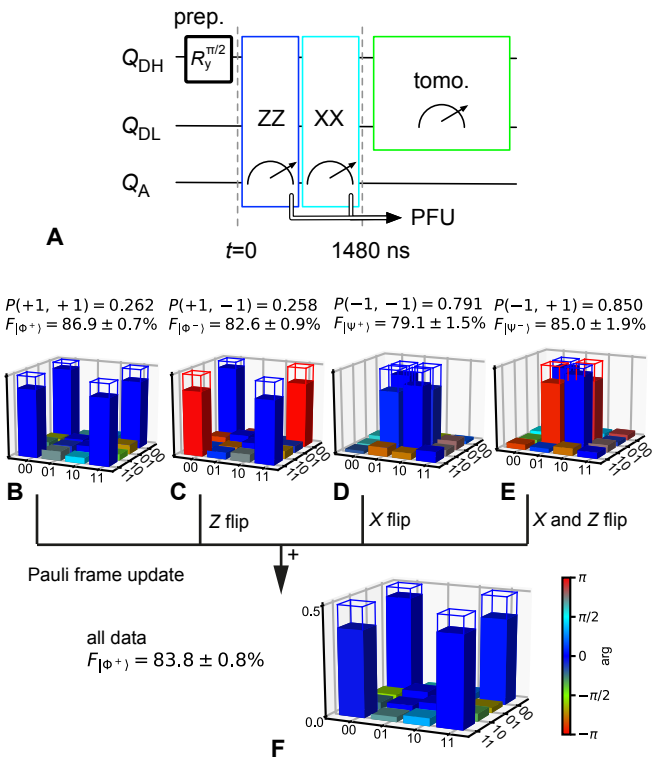

Fig. S6. Generating entanglement by sequential ZZ and XX parity measurements and PFU. (**A**) Simplified quantum circuit for preparation, ZZ and XX measurements, sequential data-qubit state tomography and PFU. (**B** to **E**) Manhattan-style plots of the reconstructed data-qubit density matrix conditioned on the ancilla measurement outcomes with occurrence and fidelity to the four expected Bell states. (**F**) We use the two-bit outcome of the parity checks to apply a PFU that transforms all runs ideally to $|\Phi^+\rangle$. Frames on the tomograms indicate the Bell states ideally produced.

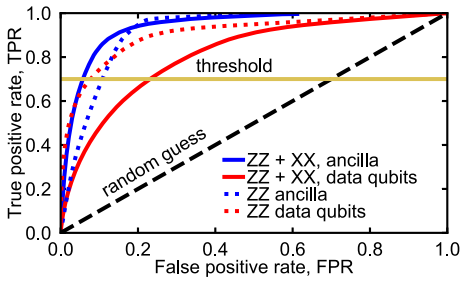

Fig. S7. ROCs for mitigation of data-qubit and ancilla leakage during interleaved ZZ and XX checks. Data-qubit and ancilla leakage are each discerned via a dedicated HMM (full curves). For comparison, the ROCs for the HMMs for repeated ZZ checks only are also shown (dotted curves, same data as in Fig. 3F).

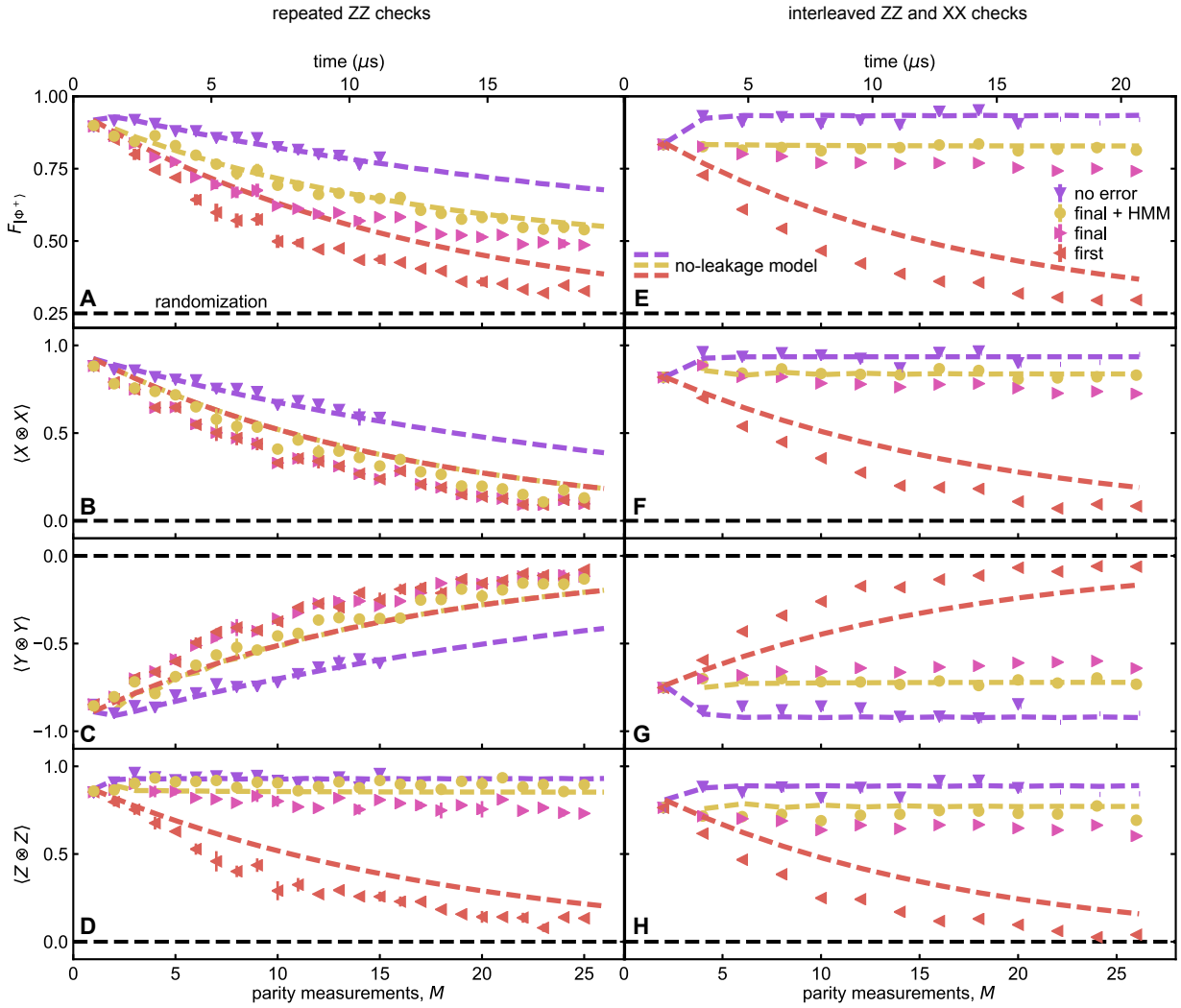

Fig. S8. Comparison of experimental data and no-leakage modeling of the repeated parity check experiments of Figs. 2 and 4. Simulations use the independently measured T_2^{echo}, T_1, e_{a}, e_{SQ}, e_{CZ}) of Table S1. This modeling uses two-level systems (no leakage) following Ref. (*36*), which uses quantumsim (*52*). As expected, the modeling is outperforming the experiment for 'first' and 'final' correction strategies as the modeling does not include leakage. It however shows an excellent matching for the 'no error' conditioning (which rejects both qubit errors and leakage). The 'final + HMM' is excellently matching the 'final' modeling curve, confirming the leakage detection capability of the HMMs.

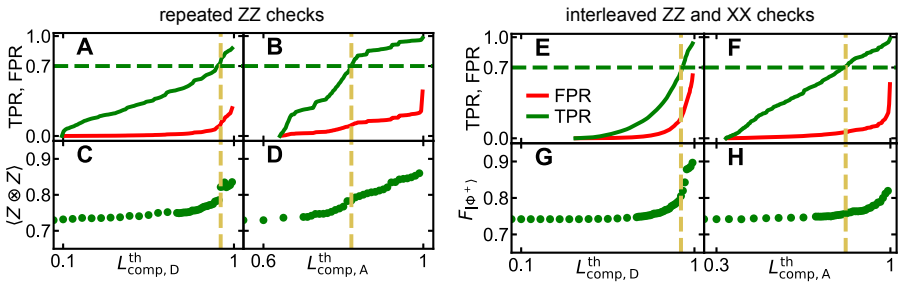

Fig. S9. Leakage mitigation for the repeated parity check experiments as a function of the chosen threshold. (**A**) [(**B**)] TPR, FPR as a function of the chosen computational-space likelihood threshold for the repeated parity check experiments of Figs. 2 and 3 for data-qubit leakage [ancilla leakage] at $M = 25$. (**C**) [(**D**)] The improvement in repeated ZZ checks is expressed as the increase in $\langle Z \otimes Z \rangle$ for data-qubit leakage [ancilla leakage]. Horizontal dashed lines indicate the chosen threshold $\mathrm{TPR} = 0.7$ (Fig. 3, F and G) and vertical dashed lines indicate the accompanying computational-space likelihoods. (**E**-**H**) Similar plots for leakage rejection for interleaved ZZ and XX checks (Fig. 4) at $M = 26$. The protocol improvement is here expressed as an increase of $F_{|\Phi^+\rangle}$.

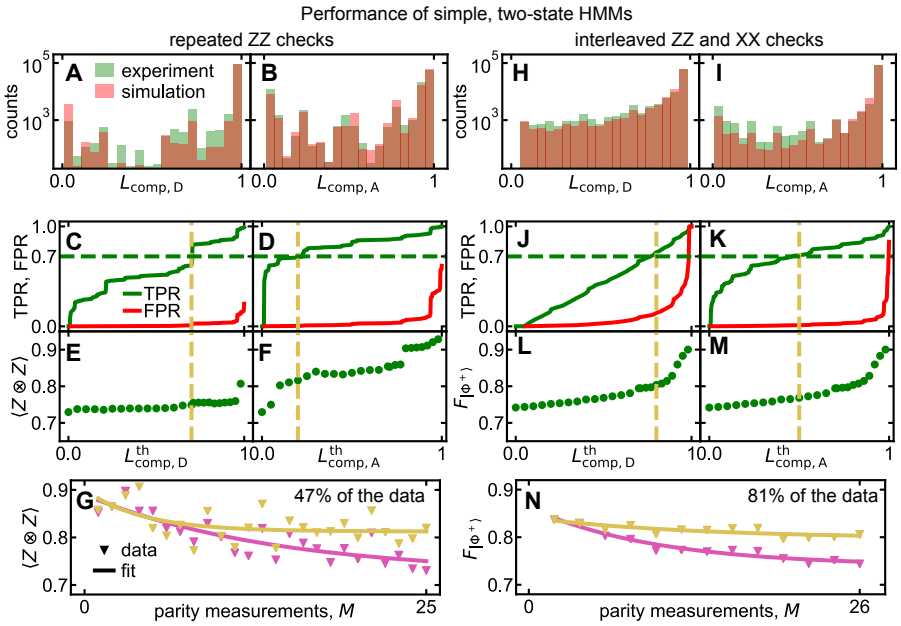

Fig. S10. Leakage mitigation for the simple, two-state HMMs for repeated parity check experiments as a function of the chosen threshold. (**A**) [(**B**)] Histograms of 10^5 $\vec{s}$ with $M = 25$ for repeated ZZ checks (as in Fig. 3D [Fig. 3E]). HMM training suggests 3.6% [20%] total data-qubit [ancilla] leakage at $M = 25$. (**C**) [(**D**)] TPR, FPR as a function of the chosen computational-space likelihood threshold for the repeated parity check experiments of Fig. 2 for data-qubit leakage [ancilla leakage] at $M = 25$. (**E**) [(**F**)] The improvement in repeated ZZ checks is expressed as the increase in $\langle Z \otimes Z \rangle$ for data-qubit leakage [ancilla leakage]. Horizontal dashed lines indicate the chosen threshold $\text{TPR} = 0.7$ and vertical dashed lines indicate the accompanying computational-space likelihoods (as in Fig. S9). (**G**) $\langle Z \otimes Z \rangle$ after M ZZ checks and correction based on the 'final' outcomes, without (same data as in Fig. 2D) and with leakage mitigation by postselection ($\text{TPR} = 0.7$). (**H**-**M**) Similar plots for simple-HMM leakage rejection for interleaved ZZ and XX checks (Fig. 4) at $M = 26$. (**N**) $F_{|\Phi^+\rangle}$ after M interleaved checks and correction based on the 'final' outcomes, without (same data as in Fig. 4) and with leakage mitigation by postselection ($\text{TPR} = 0.7$). The protocol improvement (L, M and N) is here expressed as an increase of $F_{|\Phi^+\rangle}$.

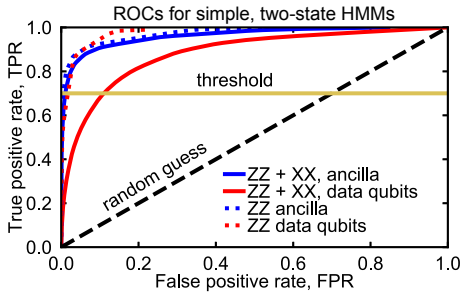

Fig. S11. ROCs for leakage mitigation as in fig. S7 but using simple two-state HMMs.