Quantifying error and leakage in an encoded Si/SiGe triple-dot qubit

R. W. Andrews, C. Jones, M. D. Reed, A. M. Jones, S. D. Ha, M. P. Jura, J. Kerckhoff, M. Levendorf, S. Meenehan, S. T. Merkel, A. Smith, B. Sun, A. J. Weinstein, M. T. Rakher, T. D. Ladd, and M. G. Borselli
HRL Laboratories, LLC, 3011 Malibu Canyon Rd., Malibu, CA, 90265, USA

Quantum computation requires qubits that satisfy often-conflicting criteria, including scalable control and long-lasting coherence [1]. One approach to creating a suitable qubit is to operate in an encoded subspace of several physical qubits. Though such encoded qubits may be particularly susceptible to leakage out of their computational subspace, they can be insensitive to certain noise processes [2, 3] and can also allow logical control with a single type of entangling interaction [4] while maintaining favorable features of the underlying physical system. Here we demonstrate a qubit encoded in a subsystem of three coupled electron spins confined in gated, isotopically enhanced silicon quantum dots [4, 5]. Using a modified "blind" randomized benchmarking protocol that determines both computational and leakage errors [6, 7], we show that unitary operations have an average total error of 0.35%, with 0.17% of that coming from leakage driven by interactions with substrate nuclear spins. This demonstration utilizes only the voltage-controlled exchange interaction for qubit manipulation and highlights the operational benefits of encoded subsystems, heralding the realization of high-quality encoded multi-qubit operations [4, 8].

Electrons trapped in silicon heterostructures have many attractive features, including very long coherence times in isotopically enriched material [9, 10] and compatibility with standard fabrication techniques. Single-spin qubits have recently demonstrated high-fidelity RF-controlled single-qubit operations [10, 11] and two-qubit gates using the exchange interaction [12–14]. However, using RF signals for single-qubit control requires a large, stable magnetic field and introduces challenges with crosstalk. Fortunately, electron spins are particularly well-suited to forming encoded qubits. Two coupled electron spins can be operated at near-zero magnetic field as a “singlet-triplet” qubit [15, 16]. That qubit is insensitive to uniform magnetic field fluctuations but still requires a magnetic field gradient for universal control. Three coupled electrons [17] can form a qubit with a tunable electric dipole moment, which could enhance RF selectivity, or the exchange-only qubit, which can be universally controlled using only the exchange interaction and does not require synchronization of gate operations with a local oscillator. Exchange is highly local and can be accurately controlled with a large on-off ratio using only fast voltage pulses. The combination of these features makes the exchange-only qubit especially attractive for use in many-qubit systems.

Here we demonstrate high-fidelity operation of a triple-dot, exchange-only qubit. Electrons are trapped in quantum dots formed by the electrostatic potentials of a Si/SiGe heterostructure and the patterned metallic gates shown in Fig. 1a [5]. As depicted in Fig. 1b, we manipulate the chemical potential of the dots by primarily adjusting voltages applied to gates P1, P2, and P3 and the exchange coupling between neighboring electrons by adjusting voltages applied to gates X1 and X2. The three coupled electrons constitute an 8-dimensional spin sys-
tem that can be used to form an exchange-only qubit [4, 17].

Operation of this encoded qubit is conveniently described in terms of the coupled spin states $|S_{12}, S, m\rangle$, where $S$ is total spin quantum number, $S_{12}$ is the total spin of the electrons confined beneath gates P1 and P2, and $m$ is the spin projection in an arbitrary direction, such as that of an applied magnetic field. The 8 possible states are enumerated in Fig. 1c. We use energy-selective initialization [5, 16] to prepare the first two electrons in a spin singlet state, giving $S_{12}=0$. Since the spin state of the third electron remains random, the qubit is initialized in an incoherent mixture of $|0, \frac{1}{2}, \frac{1}{2}\rangle$ and $|0, \frac{1}{2}, -\frac{1}{2}\rangle$. These states collectively comprise encoded 0 (denoted hereafter as $|0\rangle$). In this encoding, the quantum number $m$ does not affect logical operations and the qubit is insensitive to global magnetic fields. We measure the $S_{12}$ quantum number (which distinguishes between $|0\rangle$ and the six other spin states) via spin-to-charge conversion [5, 16] and charge-state detection using a dot charge sensor (formed using gates whose leads are visible in the top of Fig. 1a) and associated amplification circuitry [5, 18]. We denote $P_0$ as the projector onto $S_{12} = 0$ and the probability of being found in $|0\rangle$ for a density matrix $\rho$ as $\text{Tr}(P_0\rho)$. More detail on this subsystem algebra appear in the Supplemental Information.

To manipulate the encoded qubit state, we induce interaction between neighboring pairs of electrons by applying voltage pulses to device gates in order to lower the inter-dot potential barrier (see Fig. 1b). The resulting electron-electron exchange interaction only affects the $S_{12}$ quantum number [4, 5, 19], so only states within the $S = 1/2$ subspace are accessible through exchange. However, other physical mechanisms may alter the total spin and drive the qubit into leakt states with $S = 3/2$; in our case, leakage is generally caused by magnetic gradients such as those due to hyperfine interactions with nuclear spins. In the Bloch sphere representation, exchange between the two pairs of electrons induces rotations about non-orthogonal axes separated by 120 degrees [4, 17, 20]. As diagrammed in Fig. 1d, we label them as $\hat{z}$ and $\hat{n}$.

To demonstrate high-fidelity encoded operations, we first create a mapping between the voltage of the pulses we apply and the resulting exchange rotation. A rotation about the $\hat{n}$ axis involves voltage pulses on the P2, X2, and P3 gates, labelled in Fig. 1b. We choose the ratios of the voltages applied to each of these gates so that we primarily alter the potential barrier between the electrons while minimizing differences in the chemical potentials of the dots (see Fig. 2a). This “symmetric operation” minimizes sensitivity to charge noise [21, 22]. The relative gate voltages required for symmetric operation can be found using a procedure outlined in [21]. For $\hat{n}$ rotations in this device, we pulse the voltage along the vector $V(V_n) = V_0 + V_n \{-0.20, 1, -0.25\}$ in the (P2, X2, P3) voltage coordinates. Between pulses, gates are biased at an “idle” configuration $V_0$ in which all electrons are isolated and exchange rates are negligible. We map out qubit rotation by initializing the qubit, applying a sequence of three consecutive voltage pulses that are 10 ns wide and separated by an idle time of 20 ns (diagrammed in Fig. 2a), and measuring the resulting probability of finding the qubit in $|0\rangle$ ($S_{12}=0$) after applied voltage pulses. Voltage pulses cause the qubit to rotate about the $\hat{n}$ axis, leading to a maximum expected contrast of 0.75 (indicated by dashed lines).  

Calibrated rotations about the two axes can be interleaved to generate arbitrary single-qubit operations [23]. For example, an encoded Hadamard gate is applied with a sequence of a 0.96 radian rotation about the $\hat{z}$ axis, a 4.37 radian rotation about the $\hat{n}$ axis, and then a second 0.96 rotation about the $\hat{z}$ axis; this type of sequence is diagrammed in Fig 3b. For the construction used here, the 24 quantum gates in the Clifford group require an average of 2.7 individual rotations about some combination of the $\hat{n}$ and $\hat{z}$ axes, with the shortest Clifford requiring one rotation and the longest four. 

---

**FIG. 2.** Voltage-to-angle calibration of the $\hat{n}$ axis. a, To perform our calibration sequence, the qubit is initialized in $|0\rangle$ and then three voltage pulses are applied to the P2, X2, and P3 gates with the relative pulse amplitudes and durations specified in the main text. Voltage pulses are drawn to highlight non-idealities arising from control bandwidth limitations. After rotation, the qubit is measured. b, Measured probability of finding the qubit in $|0\rangle$ ($S_{12}=0$) after applied voltage pulses. Voltage pulses cause the qubit to rotate about the $\hat{n}$ axis, leading to a maximum expected contrast of 0.75 (indicated by dashed lines). c, By unwrapping the accrued phase in sinusoidal waveforms like that in b, we create a map between applied voltage and rotation from a single pulse (see Supplement for details).
less efficient than possible with systems that have orthogonal rotation axes, where Clifford operations require an average of 1.9 rotations [11]. The advantage of the triple-quantum-dot, however, lies in its simplicity: all operations only require nearest-neighbor voltage-controlled exchange interactions. Two-qubit gates can also be implemented using only exchange between two neighboring triple-dot encoded qubits [4, 8].

We demonstrate high-fidelity operation of this encoded qubit using a modified form of randomized benchmarking (RB) [24, 25]. A conventional RB experiment (diagrammed in Fig. 3a and b) consists of initializing the qubit to $|0\rangle$, applying a series of $N$ gates randomly chosen from the Clifford group, applying a final “recovery” operation $C_R$ that ideally inverts the preceding sequence, and measuring the resulting $y_0(N) = \text{Tr}[P_0\rho(N)]$, where $\rho(N)$ is the qubit density matrix after $N$ gates and the recovery gate. The average gate error rate $\varepsilon$ is determined by sweeping $N$ and fitting $A + B(1 - 2\varepsilon)^N$ to $y_0$. Data resulting from such an experiment are shown as blue data points in Fig. 3c. These data reveal that leakage plays a significant role in qubit operation: The $|0\rangle$ probability approaches 0.25 after many Clifford operations, indicating that the final state is a uniform mixture of the $8$ three-electron spin states.

We extend conventional RB to measure both qubit errors and leakage errors by exploiting two features of our qubit. First, all leaked states have $S_{12} = 1$ and so yield the same measurement outcome as $|1\rangle$. Second, exchange operations preserve $S$ and do not cause leakage. For each RB sequence, we randomly select the recovery operation to either return the qubit to $|0\rangle$ (as in conventional RB) or $|1\rangle$, keeping track of which recovery was used while still making the experiment “blind” to the correct final state. Similar modifications to the recovery operation have been studied in other leakage-detecting variants of RB [6, 7, 26, 27]. The measurement probabilities as a function of $N$ are binned in two groups $y_0$ and $y_1$, where the subscript denotes the selection of recovery operation, as shown in Fig. 3c. We extract error and leakage using the ansatz (see Supplemental Information for details)

$$y_0 = A + B(1 - p)^N + C(1 - q)^N, \quad y_1 = A - B(1 - p)^N + C(1 - q)^N. \quad (1)$$

The problem of fitting a coupled system of equations is made simpler by separately fitting single exponentials to $y_0 + y_1$ and $y_0 - y_1$ to estimate parameters $A, B, C, p$ and $q$ (see Fig. 3d).

To define error and leakage per gate, we must re-examine the effect of state preparation and measurement (SPAM) errors. In conventional RB, the dimensionality of the system is an assumption (e.g. $d = 2$ for one qubit) that becomes a normalization factor in determining gate error. With leakage, the situation is more complicated because the number of accessible leaked states might not be known. Moreover, if there are multiple leaked states, then population transfer to these states may occur at different rates since only the qubit subsystem undergoes Clifford twirling to depolarize errors. To address these complications, we define the leakage error $\Gamma$ and total error $\varepsilon$ using the decay rates $p$ and $q$, as well as parameters $B$ and $C$ that are typically associated with SPAM (see Supplemental Information for complete derivation). We assign total error to all of the population leaving subspace $P_0$ in the first gate, normalized by SPAM fidelity $2B$:

$$\varepsilon = \text{Tr} \{ P_0[\rho(0) - \rho(1)] \} / (2B) = p/2 + Cq/(2B). \quad (3)$$

If there is no leakage, $C = 0$ and $\varepsilon$ coincides with the gate error from conventional RB. The leakage probability is likewise defined as the total population leaking from

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig3}
\caption{Randomized benchmarking of an encoded qubit. a, Benchmarking involves initializing the qubit ($\rho = |0\rangle\langle 0|)$, applying a sequence of $N$ gates drawn randomly from the Clifford group ($C_i^N$), applying a recovery operation ($C_R$), and measuring the result. b, Example pulse sequence for $N = 2$. Qubit control is performed with interleaved voltage pulses that induce rotations about the $\hat{n}$ and $\hat{z}$ axes. c, $|0\rangle$ probability versus sequence length. For each sequence, the recovery Clifford is randomly selected to ideally return the state to either $|0\rangle$ ($y_0$ data shown in blue) or $|1\rangle$ ($y_1$ data in red). Error bars indicate the standard error of the mean probability from 60 sequence repetitions, and when not visible are smaller than the size of data points (see Methods). Vertical arrows label curves denoting the estimated population in states $|0\rangle$ (coinciding with $y_0$ fit), $|1\rangle$ (the shaded area between $y_0$ and the curve above), and $|Q\rangle$ (the remaining population) as a function of $N$. See Methods for additional data acquisition details. d, Single-exponential fits to sum and difference data traces.}
\end{figure}
the encoded qubit subspace after the first gate, again normalized by $2B$:

$$\Gamma = \text{Tr}\{|P_L[\rho(0) - \rho(1)]/(2B) = Cq/B, \tag{4}$$

where $P_L$ is the projector onto leaked states. Applying this analysis to the data depicted in Fig. 3c, we obtain a total error rate $\varepsilon = 0.35\%$, where leakage $\Gamma = 0.17\%$ contributes approximately half of the error. We identify the non-leak error ($\varepsilon - \Gamma$) as the “qubit error.” We also extract a SPAM fidelity $F_{\text{SPAM}} = 0.5 + B = 99.2\%$.

To confirm that blind RB correctly extracts leakage, we introduce intentional overrotation errors on the control pulses. Exchange preserves total spin $S = 1/2$, and so control imprecision only causes qubit errors and not leakage. We modify each value $\theta_i$ in the set of angles used to construct the Clifford operations as $\theta_i \rightarrow (1 + \Delta)\theta_i$, where the sign of overrotation is randomized for each angle and $\Delta$ is fixed. The gate voltages necessary to produce these rotations are found using calibrations like the one shown in Fig. 2c. We perform blind RB for values of $\Delta$ from 0%-6%, with each value repeated several times with different random overrotation signs. The resulting leakage and total error rates are shown in Fig. 4c. As expected, the total error increases with $\Delta$ but leakage remains constant at $\Gamma \approx 0.15\%$. We also see that blind RB reliably differentiates between qubit error and leakage error: for $\Delta = 6\%$, shown in Fig. 4a and 4b, we resolve a leakage error per Clifford of only 0.15% from a total error of 2.0%, a difference which leads to a distinctly non-exponential decay in $y_0$.

We use a combination of blind RB and spin simulations to test our understanding of the error per Clifford gate. We perform blind RB as a function of the idle time ($t_{\text{idle}}$) since hyperfine interactions are anticipated to increase both total error and leakage quadratically with $t_{\text{pulse}} + t_{\text{idle}}$ [20]. Fig. 4 shows that for $t_{\text{idle}} > 30$ ns, error increases as expected and the ratio of leakage to total error approaches 0.5. However, for shorter idle times, qubit error increases relative to leakage. The high-$t_{\text{idle}}$ behavior is well-replicated by detailed simulations of three spins subject to both charge and magnetic noise. Charge noise is modeled as fluctuating exchange with noise power similar to that reported in Ref. 21 and magnetic noise as independently fluctuating vector magnetic fields at each dot consistent with contact hyperfine coupling to the 800 ppm $^{29}$Si and natural $^{73}$Ge nuclear environment. Both noise sources are modeled with a $1/f$ noise spectral density [5] (see the Supplementary Information). We suspect the divergence between the simulations and the data at short $t_{\text{idle}}$ is due to the finite bandwidth of our voltage pulses preventing accurate rotations. Such control errors are similar in nature to overrotation error, do not cause leakage, and become worse with shorter $t_{\text{idle}}$ due to unintentional pulse overlap. Using the same expression relating overrotation to qubit error (Fig. 4c caption), the combined effects of imperfect control and charge noise can be modeled as stochastic overrotation. For total error $\varepsilon = 0.35\%$ and leakage $\Gamma = 0.17\%$ at the minimum total error point of $t_{\text{idle}} = 20$ ns, the root-mean-square overrotation amount is about 1.3%.

Despite the necessity of multi-pulse composite rotations and its susceptibility to leakage out of the encoded subspace, we have shown that our implementation of an exchange-only qubit using isotopically-enhanced silicon is capable of thousands of coherent qubit manipulations. This demonstration, performed at a low magnetic field and requiring only baseband voltage control, bodes well for future demonstrations with larger quantum dot arrays. Improvements in qubit error are still required, however. Through its isolation of leakage error, the blind randomized benchmarking protocol shows that total error is primarily limited by a combination of dephasing from
nuclear spins and control error. While various strategies for dynamically correcting for magnetic gradients are known [28], the most straightforward path for reducing nuclear dephasing is faster pulsing. Reducing the effects of control errors, particularly when operating faster, will require improvements of control bandwidth, software-based precompilation of pulses to accelerate pulse settling [29], or a combination thereof. These engineering concerns will be of particular importance when tackling the longer pulse sequences required for two-qubit exchange-only gates [4, 8] and beyond.

METHODS

The device utilized here was fabricated with an overlapping aluminum gate architecture similar to previously published devices [21, 30, 31]. The device was measured in a dilution refrigerator with a base temperature of 20 mK. High bandwidth coaxial cables allowed for rapid pulsing of the voltages applied to gates P1, X1, P2, X2, P3, necessary for performing randomized benchmarking at the point of symmetric operation. Voltage pulses were sourced by arbitrary waveform generators with a 400 MHz sample rate, the outputs of which were low-pass filtered with a cutoff of 120 MHz and a flat time delay. All experiments were performed with 10 ns pulse time and 20 ns idle time (unless indicated otherwise) and with zero applied magnetic field.

Blind randomized benchmarking was performed by generating a random sequence of Clifford gates of a given length, \( N \), with the final Clifford randomly chosen to invert to \(|0\rangle\) or \(|1\rangle\). This sequence is combined with state preparation and measurement and then averaged over 100 runs before moving to the next sequence length. The length of the Clifford sequence is then swept sixty times, from \( N = 1 \) to \( N = 2000 \). The sixty measurements for each sequence length are combined, and their average and standard deviation are the data points and error bars shown in Fig. 3c. The same averaging conditions were used for the data of Fig. 4.

Measurement of \( S_{12} \) was accomplished through spin-to-charge conversion and charge sensing, with high SNR (~9) single-shot discrimination enabled by cryogenic HEMT amplification, as in Ref. 18. Each measurement was individually thresholded to map each result to 0 or 1. To avoid populating excited states when transitioning between gate voltages used for measurement and idle configurations, voltages on gates P1 and P2 were slewed adiabatically when near the region of spin-to-charge conversion. The total sequence length of state preparation, adiabatic ramping, and measurement with a single exchange pulse was 137.1 \( \mu s \).

[1] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien, “Quantum computers,” Nature 464, 45 (2010).
[2] P. Zanardi and M. Rasetti, “Noiseless quantum codes,” Phys. Rev. Lett. 79, 3306 (1997).
[3] D. A. Lidar, I. L. Chuang, and K. B. Whaley, “Decoherence-free subspaces for quantum computation,” Phys. Rev. Lett. 81, 2594 (1998).
[4] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley, “Universal quantum computation with the exchange interaction,” Nature 408, 339 (2000).
[5] K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, “Isotopically enhanced triple-quantum-dot qubit,” Sci. Adv. 1, e1500214 (2015).
[6] J. J. Wallman, M. Barnhill, and J. Emerson, “Robust characterization of leakage errors,” New J. Phys. 18, 043021 (2016).
[7] C. J. Wood and J. M. Gambetta, “Quantification and characterization of leakage errors,” Phys. Rev. A 97, 032306 (2018).
[8] B. H. Fong and S. M. Wandzura, “Universal quantum computation and leakage reduction in the 3-qubit decoherence free subsystem,” Quantum Info. Comput. 11, 1003 (2011).
[9] A. M. Tyryshkin, S. Tojo, J. J. L. Morton, H. Riemann, N. V. Abrosimov, P. Becker, H.-J. Pohl, T. Schenkel, M. L. W. Thewalt, K. M. Itoh, and S. A. Lyon, “Electron spin coherence exceeding seconds in high-purity silicon,” Nat. Mater. 11, 143 (2011).
[10] Z. T. Yoneda, K. Takeda, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tanaka, “A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%,” Nat. Nanotechnol. 13, 102 (2018).
[11] E. Kawakami, T. Jullien, P. Scarlino, D. R. Ward, D. E. Savage, M. G. Legally, V. V. Dobrovitski, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, “Gate fidelity and coherence of an electron spin in an Si/SiGe quantum dot with micromagnet,” Proc. Natl. Acad. Sci. USA 113, 11738 (2016).
[12] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta, “Resonantly driven CNOT gate for electron spins,” Science 359, 439 (2017).
[13] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Legally, Mark Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, “A programmable two-qubit quantum processor in silicon,” Nature 555, 633 (2018).
[14] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, A. Morello, A. Laucht, and A. S. Dzurak, “Fidelity benchmarks for two-qubit gates in silicon,” arXiv:1805.05027.
[15] J. Levy, “Universal quantum computation with spin-1/2 pairs and Heisenberg exchange,” Phys. Rev. Lett. 89, 147902 (2002).
[16] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossett, “Coherent manipulation of coupled electron spins in semiconductor quantum dots,” Science 309, 2180 (2005).

[17] M. Russ and G. Burkard, “Three-electron spin qubits,” J. Phys. Condens. Matter 29, 393001 (2017).

[18] A. M. Jones, E. J. Pritchett, E. H. Chen, T. E. Keating, R. W. Andrews, J. Z. Blumoff, L. A. De Lorenzo, K. Eng, S. D. Ha, A. A. Kiselev, S. M. Meenehan, S. T. Merkel, J. A. Wright, L. F. Edge, R. S. Ross, M. T. Rakher, M. G. Borselli, and A. Hunter, “Spin-blockade spectroscopy of Si/SiGe quantum dots,” arXiv:1809.08320.

[19] D. Loss and D. P. DiVincenzo, “Quantum computation with quantum dots,” Phys. Rev. A 57, 120 (1998).

[20] T. D. Ladd, “Hyperfine-induced decay in triple quantum dots,” Phys. Rev. B 86, 125408 (2012).

[21] M. D. Reed, B. M. Maune, R. W. Andrews, M. G. Borselli, K. Eng, M. P. Jura, A. A. Kiselev, T. D. Ladd, S. T. Merkel, I. Milosavljevic, E. J. Pritchett, M. T. Rakher, R. S. Ross, A. E. Schmitz, A. Smith, J. A. Wright, M. F. Gyure, and A. T. Hunter, “Reduced sensitivity to charge noise in semiconductor spin qubits via symmetric operation,” Phys. Rev. Lett. 116, 110402 (2016).

[22] F. Martins, F. K. Malinowski, P. D. Nissen, E. Barnes, S. Fallahi, G. C. Gardner, M. J. Manfra, C. M. Marcus, and F. Kuemmeth, “Noise suppression using symmetric exchange gates in spin qubits,” Phys. Rev. Lett. 116, 116801 (2016).

[23] F. Lowenthal, “Uniform finite generation of the rotation group,” Rocky Mountain J. Math. 1, 575 (1971).

[24] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, “Symmetrized characterization of noisy quantum processes,” Science 317, 1893 (2007).

[25] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, “Randomized benchmarking of quantum gates,” Phys. Rev. A 77, 012307 (2008).

[26] J. Helsen, J. J. Wallman, S. T. Flammia, and S. Wehner, “Multi-qubit randomized benchmarking using few samples,” arXiv:1701.04299v2.

[27] X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and L. M. K. Vandersypen, “Benchmarking gate fidelities in a Si/SiGe two-qubit device,” arXiv:1811.04002.

[28] G. T. Hickman, X. Wang, J. P. Kestner, and S. Das Sarma, “Dynamically corrected gates for an exchange-only qubit,” Phys. Rev. B 88, 161303 (2013).

[29] B. R. Johnson, M. D. Reed, A. A. Houck, D. I. Schuster, Lev S. Bishop, E. Ginossar, J. M. Gambetta, L. DiCarlo, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, “Quantum non-demolition detection of single microwave photons in a circuit,” Nat. Phys. 6, 663 (2010).

[30] M. G. Borselli, K. Eng, R. S. Ross, T. M. Hazard, K. S. Holabird, B. Huang, A. A. Kiselev, P. W. Deelman, L. D. Warren, I. Milosavljevic, A. E. Schmitz, M. Sokolich, M. F. Gyure, and A. T. Hunter, “Undoped accumulation-mode Si/SiGe quantum dots,” Nanotechnology 26, 375202 (2015).

[31] D. M. Zajac, T. M. Hazard, X. Mi, K. Wang, and J. R. Petta, “A reconfigurable gate architecture for Si/SiGe quantum dots,” Appl. Phys. Lett. 106, 223507 (2015).

AUTHOR CONTRIBUTIONS

The device was designed, characterized, and fabricated by S.D.H., M.P.J., M.L., M.T.R., and M.G.B. Control theory, coding, and analysis were provided by C.J., A.M.J., S.M., S.T.M., A.S., A.J.W., and T.D.L. Measurements were made by R.W.A., M.D.R., A.M.J., J.K., S.M., B.S., and A.J.W. The manuscript was written by R.W.A., C.J., M.D.R., A.M.J, and T.D.L. with input from all authors. The effort was supervised by M.T.R., T.D.L., and M.G.B.

ACKNOWLEDGEMENTS

We thank Andy Hunter, Kevin Eng, Mark Gyure, and Bryan Fong for valuable contributions leading to this work.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

COMPETING INTERESTS

The authors declare no competing interests.
I. ENCODING DETAILS FOR THE EXCHANGE-ONLY QUBIT

In this section we further explain the algebra of exchange-only control of triple-dot spin qubits. This qubit is defined in a subsystem of angular-momentum states whose energy differences are immune to fluctuations in global magnetic fields [1]. If we notate the total spin of the three electrons as $S = S_1 + S_2 + S_3$ and the total spin of the first two electrons as $S_{12} = S_1 + S_2$, then by standard rules for the addition of angular momentum, all 8 spin-states may be written as $|S_{12}, S; m⟩$, where $S_{12}$ takes on the values 0 or 1 and $S$ takes on the values 1/2 and 3/2. Our qubit states are defined as

$$
|0⟩ \otimes |m⟩_{1/2} = |0, 1/2; m⟩,
|1⟩ \otimes |m⟩_{1/2} = |1, 1/2; m⟩,
|Q⟩ \otimes |m⟩_{3/2} = |1, 3/2; m⟩.
$$

Although $m$ is a collective property of all three spins, we treat it as a separate subsystem which we do not initialize, measure, or control. Critically, $m$ is the only degree of freedom coupled to global magnetic fields, and so ignoring $m$ is tantamount to ignoring all global magnetic fields, including global fields which may be fluctuating in any vector direction\(^1\). Our initial state may be written as $|0⟩|0⟩ \otimes ρ_m$, where $ρ_m$ is a fully mixed (that is, completely unknown) state for the $m$ degree of freedom. Likewise, the singlet measurement operator is the projector

$$
P_0 = |0⟩|0⟩ \otimes I_m = \sum_{m=±1/2} |0, 1/2; m⟩⟨0, 1/2; m|.
$$

The matrix elements for Heisenberg exchange between any two spins in this basis can be expressed via the Wigner-Eckart theorem [2], which first of all indicates that exchange leaves the quantum numbers $S$ and $m$ invariant, second that $(S_{12}, 3/2; m | S_j, S_k | S_{12}, 3/2; m) = 1/4$ for any spin pair $j, k$ (and hence leaked states are not affected by exchange), and finally that the matrix elements for the qubit states in terms of Wigner $6j$ coefficients (notated $\{ \}$) and reduced matrix

$$
\begin{array}{c|c|c|c|c}
\text{Gate} & \text{R}_z & \text{R}_m & \text{R}_z & \text{R}_m \\
\hline
I & - & - & - & - \\
X & \pi - \theta_1 & \theta_1 & \pi - \theta_1 & \\
Y & \pi - \theta_1 & \theta_1 & \pi - \theta_1 & \\
Z & \pi & - & - & - \\
S & 3\pi/2 & - & - & - \\
S_1 & \pi/2 & - & - & - \\
SX & - & \theta_3 & \theta_2 & \theta_4 \\
S_1X & - & \theta_4 & \theta_2 & \theta_3 \\
H & (\pi - \theta_1)/2 & \pi + \theta_1 & (\pi - \theta_1)/2 & - \\
XH & (\pi + \theta_1)/2 & \pi - \theta_1 & (3\pi + \theta_1)/2 & - \\
YH & (\pi + \theta_1)/2 & \pi - \theta_1 & (\pi + \theta_1)/2 & - \\
ZH & (3\pi + \theta_1)/2 & \pi - \theta_1 & (\pi + \theta_1)/2 & - \\
SH & (\pi - \theta_1)/2 & \pi + \theta_1 & 2\pi - \theta_1/2 & - \\
HS & 2\pi - \theta_1/2 & \pi + \theta_1 & \pi - \theta_1/2 & - \\
S_1HS & (3\pi + \theta_1)/2 & \pi - \theta_1 & \theta_1/2 & - \\
HS_1 & \theta_1/2 & \pi - \theta_1 & (3\pi + \theta_1)/2 & - \\
S_1SH & \theta_1/2 & \pi - \theta_1 & \theta_1/2 & - \\
S_1XH & (\pi + \theta_1)/2 & \pi - \theta_1 & \theta_1/2 & - \\
HXS & 2\pi - \theta_1/2 & \pi + \theta_1 & (3\pi - \theta_1)/2 & - \\
SXS & (3\pi - \theta_1)/2 & \pi + \theta_1 & 2\pi - \theta_1/2 & -
\end{array}
$$

\(^1\) Note that the two encoded states $|0⟩$ and $|1⟩$ are each associated with two values of $m = ±1/2$ and the leaked state $|Q⟩$ is associated with four $|m⟩$ states $m = -3/2, \ldots, 3/2$; here $Q$ stands for quadruplet. The varying number of $|m⟩$ states means the subsystem separation is imperfect; a true tensor product of subsystems could be recovered by treating the $|m⟩$ subspace as four-dimensional and imposing a selection rule preventing population of $|S_{12}, 1/2; ±3/2⟩$ states.
elements (notated $\langle || \cdot || \rangle$) are given by

\[
\langle S_{12}, 1/2; m | S_1 \cdot S_2 | S'_{12}, 1/2; m \rangle = \delta_{S_12S'_12} (-1)^{S_12+1/2} \left\{ \begin{array}{ccc} 1/2 & 1/2 & 1/2 \\ S_12 & 1 & 2/2 \end{array} \right\} (1/2||S||1/2)^2
\]

\[
= -\frac{1}{4} - \frac{1}{2} \sigma^z,
\]

(3)

\[
\langle S_{12}, 1/2; m | S_2 \cdot S_3 | S'_{12}, 1/2; m \rangle = (-1)^{S_12+S'_12} \sqrt{(2S_{12} + 1)(2S'_{12} + 1)} \left\{ \begin{array}{ccc} S_12 & 1 & S'_{12} \\ 1/2 & 1/2 & 1/2 \end{array} \right\}^2 (1/2||S||1/2)^2
\]

\[
= -\frac{1}{4} + \frac{1}{4} \sigma^z + \frac{\sqrt{3}}{4} \sigma^x,
\]

(4)

where $\sigma^x$ and $\sigma^z$ are $2 \times 2$ Pauli matrices in the basis of the $|0\rangle$ and $|1\rangle$ states. Neglecting overall phase, we therefore define the effective exchange-only Hamiltonian

\[
H_{\text{exchange}}(t) = J_z(t) S_1 \cdot S_2 + J_n(t) S_2 \cdot S_3 \rightarrow \frac{1}{2} J_z(t) \sigma^z - \frac{1}{2} J_n(t) [\sigma^x \cos \phi - \sigma^x \sin \phi],
\]

(5)

where $\phi = 2\pi/3$. Hence in the $|0\rangle/|1\rangle$ Bloch sphere, exchange between dots 1 and 2 causes rotations about the $z$-axis, and exchange between dots 2 and 3 causes rotations about the $n$-axis, which is rotated from the $z$-axis in a right-handed sense about the $y$-axis by $120^\circ$.

Qubit control is accomplished using generalized Euler angles by pulsing a single axis at a time. Table S1 shows all of the angles used to generate the 24 single-qubit Cliffords. As an example, consider the Hadamard gate $H$, which is constructed via a 3-pulse sequence. The middle pulse turns on $J_z$ at a level to drive a rotation of angle $\int_0^\tau J_n(t) dt = \pi + \tan^{-1}\sqrt{8}$ in pulse time $\tau$. This angle is chosen as the one which rotates the $z$-axis to the $xy$ plane. Additional $z$-rotations convert this rotation to a Hadamard; our composite gate therefore begins and ends by pulsing $J_z$ for a total angle of $\int_0^\tau J_z(t) dt = \tan^{-1}\sqrt{2}$. The other Cliffords are constructed similarly; note that these constructions are not unique, but are guaranteed to exist and to require no more than four pulses for any single-qubit gate [3].

Local magnetic field gradients, most typically caused by hyperfine interactions with $^{28}$Si nuclear spins, cause leakage from the $|0\rangle$ and $|1\rangle$ states into the $|Q\rangle$ states. They are therefore not describable within the Bloch sphere picture. For magnetic fields sufficiently high to ignore electron-nuclear flip-flop terms of the contact hyperfine interaction, an SU(3) picture is helpful to quantify leaking during and between pulses [4]. Note that by the addition rules for angular momentum, $S_{12} = 1$ for all $S = 3/2$ leaked states, and therefore all leaked states appear as triplet states on spins 1 and 2, and are measured as such in spin-blockade.

II. DETAILS OF ROTATION CALIBRATION

The purpose of calibration is to generate a one-to-one mapping between exchange rotation and voltage throw for symmetric operation of the respective axis [5]. To formulate this mapping, we apply a two-stage procedure consisting of ‘rough’ and ‘fine’ calibration stages.

In the ‘rough’ calibration step, we estimate a functional form of $f_\theta = \theta_\sigma(V_\sigma)$ that relates applied voltage $V_\sigma$ to rotation $\theta_\sigma$ about the $\sigma = n, z$ axes. As described in Fig. 2b of the main text, we throw three consecutive exchange pulses with equal voltage amplitudes chosen to be along a vector $V$ which has been calibrated to account for capacitive cross-talk of neighboring gates. This ‘symmetric operation tangent’ primarily adjusts the height of the potential barrier between electrons while keeping their chemical potential fixed. A full rough calibration (shown in Fig. 2b of the main text) is acquired by performing numerous experiments whose amplitudes are chosen to be equally spaced along the symmetric operation tangent. We find initial estimates of coordinates $\langle V, \theta \rangle_R$ by mapping the turning points of $|0\rangle$ probability to integer multiples of $\pi$. The subscript ‘R’ indicates that this is a rough estimate of the voltage-exchange pair. We initially fit these seed coordinates to a function of the form

\[
f_\theta = V_n(\theta_{n,R}) = -2A \times \ln \left[ \frac{\theta_{\text{max}} - \theta_{n,R}}{\theta_{\text{max}} \sqrt{\theta_{n,R}}} \right] + B
\]

(6)

where $A$ and $B$ are constants in units of voltage, and $\theta_{\text{max}}$ is the asymptotic maximum of rotation angle. For the data shown in Fig. 2b of the main text, $A = 24.24$, $B = 59.29$, and $\theta_{\text{max}} = 32.96$; the resulting curve is shown in Fig. S1a.
In the ‘fine’ calibration step, we apply a series of 10 pulses at voltages that we now linearly sample in the ‘rough’ calibration domain specified by Eqn. 6. That is, we resample the voltages along the symmetric operating tangent as $V_n = f_θ(θ_i)$ where $θ_i$ is linearly sampled over the user-specified range $[θ_{min}, θ_{max}]$. The resulting narrow-band signal appears nearly monochromatic (Fig. S1b) and the corrections needed to create an accurate calibration are encoded as deviations from a single-frequency oscillation. Because of this, we can strongly reject fluctuations in the data that occur at both high and low frequencies as noise. We analyze the data by first extending it using Burg’s maximum entropy method [6], and then apply a FIR filter with a passband of 0.5 to 2 times the signal frequency.

To estimate phase, we construct the analytic signal of the filtered, smoothed signal via the Hilbert transform and calculate the instantaneous phase as

$$\theta_n(V_n) = \frac{1}{N_p} \text{ArcTan} \left[ H \left( \tilde{P}(V_n) \right), \tilde{P}(V_n) \right],$$

where ArcTan() is the four-quadrant inverse tangent, $N_p = 10$ is the number of pulses, $H$ is the Hilbert transform and $\tilde{P}(V_n)$ is the filtered and extended data. We remove the extension data introduced by Burg’s method and unwrap the angles to produce a set of voltage-exchange pairs $\{V_n, θ_n\}$ at every sampled voltage. For proper unwrapping of the angles, we sample at sufficiently low exchange rates to ensure that the first point falls within the branch $[0, 2π/N_p]$. Finally, we convert the mesh of voltage-exchange pairs into a continuous function via nonlinear interpolation with Eqn. 6,

$$V(θ) = f_θ \left[ (1 - \alpha) f_θ^{-1}(V_i) + α f_θ^{-1}(V_j) \right],$$

where $α = (θ - θ_i)/(θ_j - θ_i)$ and $\{V_i, θ_i\}, \{V_j, θ_j\}$ bound the interval containing the specified $θ$.

FIG. S1. Details of axis calibration. a, Rough mapping between rotation angle and applied voltage. b, Data used for ‘fine’ calibration showing state probability as a function of rough rotation angle ($θ_{n,R}$). Note that this signal is very nearly monochromatic.

### III. DERIVATION OF BLIND RB

Blind RB modifies the conventional RB experiment [7] by estimating the probability that the system has leaked out of the qubit subsystem. Whereas traditional RB only analyzes the population for a two-state system ($|0⟩$ and $|1⟩$), Blind RB estimates the population for a third state as well (labeled $|Q⟩$ in our exchange-only qubit). As an aside, there can be multiple leaked states, and this analysis considers their collective population. The intuition behind Blind RB can be seen in projectors: let $P_0$, $P_1$, and $P_L$ be the projectors onto $|0⟩$, $|1⟩$, and $|Q⟩$ (leaked states), respectively. Recall from the main text, and especially Fig. 1c, that $P_0$ and $P_1$ are two-dimensional because of the undetermined $m$ quantum number, while $P_L$ is four-dimensional (all four of the $S = 3/2$ states are leaked). Since $P_L$ spans all of the states outside of the qubit basis, we have the normalization $P_0 + P_1 + P_L = 1$. The spin-to-charge measurement reports the population in $|0⟩$ for any density matrix $ρ$ as $\text{Tr}(P_0 ρ)$, assuming no measurement errors. If we can also measure the population in $|1⟩$, then we can determine the leaked population as $\text{Tr}(P_L ρ) = 1 - [\text{Tr}(P_0 ρ) + \text{Tr}(P_1 ρ)]$. This intuition assumes no state-preparation-and-measurement (SPAM) errors; however, we show below how Blind RB is able to separate gate errors from SPAM errors in the same manner as conventional RB [7].

There is a straightforward way to measure $P_1$ if you have a measurement process for $P_0$ and control of the qubit: apply an $X$ gate just before measurement. For some state $ρ$, this simply leads to a measurement probability

$$\text{Tr} \left( P_0 X ρ X^\dagger \right) = \text{Tr} \left( X^\dagger P_0 X ρ \right) = \text{Tr} \left( P_1 ρ \right),$$

where $\text{ArcTan}()$ is the four-quadrant inverse tangent.
where we have used the cyclic property of the trace. However, to measure both $P_0$ and $P_1$, one must have multiple copies of $\rho$. This is possible in randomized benchmarking because $\rho$ is an ensemble over all Clifford circuits of some length $N$, which is approximated with some randomly sampled subset of such Clifford circuits. Hence we can apply the technique above by preparing many copies of $\rho(N)$, the state after a random $N$-Clifford circuit, and measuring some with $P_0$ and the others with an $X$ gate followed by $P_0$ to approximate $P_1$.

Blind RB makes one additional modification to measuring $P_1$ effectively. Since gates are not perfect, one might be concerned that any experiment that attempts to measure $P_1$ has slightly more error than measurement of $P_0$ due to any error in the added $X$ gate. Blind RB addresses this by modifying the final “recovery” Clifford gate: half of the Clifford sequences compose to identity (target final state is $|0\rangle$), and the other half compose to $X$ (target final state is $|1\rangle$). Because the $X$ gate is compiled into the recovery operation, one can measure $P_1$ in this circumstance without the additional gate error.

Having established the intuition for Blind RB, we now show more formally how the experiment is analyzed. Recall from the main text that the probabilities for measuring the target final state are binned in two groups $y_0(N)$ and $y_1(N)$, where subscript is the target state, as a function of sequence length $N$. If $\rho(N)$ is the state after $N$ Clifforeds, then the state after $N$ Clifforeds with the flipped recovery operation is approximately $X\rho(N)X$, and the measured probabilities are

$$y_0 = \text{Tr}[P_0\rho(N)],$$

$$y_1 = \text{Tr}[P_0X\rho(N)X] \approx \text{Tr}[P_1\rho(N)].$$

We can begin to analyze the results by first considering leakage. The sum combination can be expressed using the normalization as

$$W(N) = y_0(N) + y_1(N) \approx \text{Tr}[(P_0 + P_1)\rho(N)] = 1 - \text{Tr}[P_L\rho(N)],$$

where approximate equality accounts for the presence of SPAM errors. We fit the measured $W(N)$ to the ansatz

$$W(N) = 2A + 2C(1-q)^N,$$

where parameters $A$ and $C$ account for the unknown number of accessible leaked states and SPAM error. We conjecture that $W(N)$ decays as a single exponential in $N$. The intuition for this is that Clifford twirling in the qubit subsystem will interrupt any coherent driving between qubit and leaked states, emulating a $T_1$-like process. This conjecture is supported by the data in Fig. 3d of the main text showing that an exponential fit follows the data closely. More detailed theory for this matter is the subject of future work.

We can also extract errors within the qubit subsystem by fitting to the difference combination. Let $R(N) = y_0(N) - y_1(N)$:

$$R(N) \approx \text{Tr}[(P_0 - P_1)\rho(N)],$$

where again approximate equality accounts for the presence of SPAM errors. We fit this to

$$R(N) = 2B(1-p)^N.$$

Other works have discussed the advantage of the $(P_0-P_1)$ ansatz for fitting [8, 9], but we see that it is also advantageous in the presence of leakage since $P_0 - P_1 = I - (2P_1 + P_L)$. By randomizing what the final state of the experiment should be, a leakage event has a 50% chance of being correct, producing the same measurement statistics in $y_0 - y_1$ as a depolarizing error within the qubit subspace. As a remark, this is where “blind” comes from in the name of the protocol – during any experiment, the system is blind to whether the correct final answer should be $|0\rangle$ or $|1\rangle$ without knowing all of the Clifford gates in the sequence. Combined with standard RB assumptions for Clifford twirling of gate errors, $R(N)$ is expected to be a single exponential decay.

Echoing the main text, the total error per gate and leakage per gate are defined in the following way, using the analysis methods of randomized benchmarking to separate out SPAM errors [7]. Total error $\varepsilon$ is the population leaving $|0\rangle$ in first gate, normalized by the “visibility” of the experiment $(2B)$:

$$\varepsilon = \text{Tr} \{P_0[\rho(0) - \rho(1)]\} / (2B) = p/2 + Cq/(2B).$$

As was noted in the main text, this definition will coincide with the gate error from conventional RB when there is no leakage. Whereas conventional RB assumes the dimensionality of the qubit to be 2 by definition, we use the visibility $2B$ to normalize error, which is essentially the contrast between two experiments: (a) prepare $|0\rangle$, measure $P_0$ and (b) prepare $|0\rangle$, apply $X$, measure $P_0$. Ideally, $2B = 1$, but it can be less due to SPAM error. In the next paragraph, we
examine the effect of SPAM errors more deeply, which helps to explain why we say that we “normalize by SPAM”; what this analysis does is infer what the rate of population transfer would be (for gate error and for leakage) if there were no SPAM errors. One can apply the same analysis to estimate fidelity as

$$F_{\text{SPAM}} = 0.5 + B,$$  \hspace{1cm} (17)

which we derive below. Similar to total error, leakage $\Gamma$ is the population leaving the qubit subsystem in the first gate, also normalized by $2B$:

$$\Gamma = \text{Tr} \{ [P_C[\rho(0) - \rho(1)]] / (2B) = Cq/B. \hspace{1cm} (18)$$

We use the term “qubit error” (denoted here as $\epsilon$) to correspond to the probability of error per gate that is not leakage outside of the encoded qubit, which is simply $\epsilon = \epsilon - \Gamma$.

The fits also enable estimating the population in each of the states $|0\rangle$, $|1\rangle$, and $|Q\rangle$ during the course of the Blind RB experiment. These state probabilities are shown in Fig. 3c and Fig. 4a of the main text. We must also account for SPAM, so we introduce two more “states,” $M_0$ and $M_1$. These are the probabilities that the final measurement will report “0” or “1” regardless of the expected state of the triplet-dot system, indicating error during preparation, measurement, or both. As they are SPAM errors, $M_0$ and $M_1$ are not functions of the number of Clifford gates $N$, by assumption. In general, $M_0 \neq M_1$ if there is a bias to one outcome in the measurement process (this can occur due to T1 decay during measurement $[10, 11]$, among other reasons). Let the following be probabilities for actually measuring the correct state of the system, meaning no SPAM error:

$$\nu_0 \equiv \text{Pr(no SPAM error)} \text{Tr}(P_0 \rho); \hspace{1cm} (19)$$

$$\nu_1 \equiv \text{Pr(no SPAM error)} \text{Tr}(P_1 \rho); \hspace{1cm} (20)$$

$$\nu_L \equiv \text{Pr(no SPAM error)} \text{Tr}(P_L \rho). \hspace{1cm} (21)$$

Since $M_0 + M_1 = \text{Pr(SPAM error)}$, we have the normalization

$$\nu_0 + \nu_1 + \nu_L + M_0 + M_1 = 1. \hspace{1cm} (22)$$

Based on these definitions, we can relate these terms to the fits to $y_0$ and $y_1$ as follows:

$$y_0(N) = A + B(1 - p)^N + C(1 - q)^N = \nu_0(N) + M_0; \hspace{1cm} (23)$$

$$y_1(N) = A - B(1 - p)^N + C(1 - q)^N = \nu_1(N) + M_0. \hspace{1cm} (24)$$

By assumption, the qubit is ideally initialized as $|0\rangle$, so $\nu_1(0) = \nu_L(0) = 0$. Since we assume $M_0$ and $M_1$ are independent of $N$, we can solve for them at $N = 0$ and produce the following probabilities for all $N$:

$$M_0 = A - B + C; \hspace{1cm} (25)$$

$$\nu_0(N) = y_0(N) - M_0 = B \left[1 + (1 - p)^N\right] + C \left[(1 - q)^N - 1\right]; \hspace{1cm} (26)$$

$$\nu_1(N) = y_1(N) - M_0 = B \left[1 - (1 - p)^N\right] + C \left[(1 - q)^N - 1\right]; \hspace{1cm} (27)$$

$$M_1 = 1 - y_0(0) = 1 - (A + B + C); \hspace{1cm} (28)$$

$$\nu_L(N) = 1 - [M_0 + M_1 + \nu_0(N) + \nu_1(N)] = 2C \left[1 - (1 - q)^N\right]. \hspace{1cm} (29)$$

From this, we can derive SPAM fidelity as

$$F_{\text{SPAM}} = 1 - 0.5 (M_0 + M_1) = 0.5 + B, \hspace{1cm} (30)$$

where the factor $1/2$ comes from the fact that, averaging over all states, the errors where the measurement returns “0” or “1” regardless of input will be correct half of the time (this follows the convention where single-qubit RB error is half of the equivalent depolarizing-channel parameter $[7]$). In the main text, we plot an additional curve in each of Fig. 3c and Fig. 4a, which is $y_0(N) + \nu_1(N)$. As indicated by the vertical arrows and described in the caption, the area between this curve and $y_0(N)$ is the estimated probability for the qubit be in $|1\rangle$, as a function of sequence length $N$. Likewise, the area between this curve and above to $(1 - M_1)$ is $\nu_L(N)$, the probability of leakage. However, since $M_0$ and $M_1$ are each about 0.4%, we excluded them from the plots since they would be imperceptible.

Finally, we make a few remarks on the analysis. One can say that the parameter $q$ or the combination $y_0 + y_1$ measures leakage because of the proportionality $\Gamma \propto q$. However, it is not correct to say that parameter $p$, nor the combination $y_0 - y_1$, measures total error or qubit error. This can be seen in the definition of total error, which can

restated as \( \varepsilon = p/2 + \Gamma/2 \). In fact, the rate of exponential decay in \( y_0(N) - y_1(N) \) lies exactly halfway between twice the total error and twice the qubit error:

\[
p = 2\varepsilon - \Gamma = 2\varepsilon_q + \Gamma. \tag{31}
\]

The reason for this connects back to the discussion above about why Blind RB is called “blind.” The decay seen in \( R(N) = y_0(N) - y_1(N) \) accounts for every qubit error, but only half of the leak errors. Mathematically, this is captured in the expression above that \( P_0 - P_3 = 1 - (2P_1 + P_2) \). Alternatively, one can say that every qubit error will register as incorrect, while a leakage error will be correct half of the time. In the exchange-only qubit, a leaked state will always register as the Pauli blockaded state (i.e. “not \([0]\)”) when measured; when the correct final state is randomized, this faulty measurement will be right 50% of the time. In some contexts, it may indeed be the case that a leaked qubit can be treated as if it were depolarized; for example, the leakage-reduction unit (LRU) derived in Ref. [12] will turn a leaked exchange-only qubit into a non-leaked one, but its state relative to what it ideally would be is effectively depolarized. Assuming this LRU will be applied at some point, one could report the error per gate as simply \( p/2 \). However, we have chosen to take the conservative approach and count all of leakage probability in our reported total error.

We also note that the way Blind RB looks at Clifford sequences that should compose to identity or \( X \) (i.e. end in \([0]\) or \([1]\), respectively) is very similar to the method in Ref. [13], which proposes to simply omit the final inversion gate of an RB experiment. By leaving out the inversion gate, 1/6 of the sequences will end ideally as \([0]\), 1/6 will end as \([1]\), and the remaining 2/3 will end with 1/6 probability each in the \([\pm]\), \([\pm i]\) eigenstates of the \( X \) and \( Y \) operators. Measuring \( P_0 \) on this mixed state will give the same expectation as \( (y_0 + y_1)/2 \), but the variance will be higher for the same number of experimental samples due to measurement-projection noise from the \([\pm]\) and \([\pm i]\) states. Simply put, Blind RB focuses on the Clifford sequences that are more informative, giving more accurate estimates of leakage error for the same number of experimental runs.

\section*{IV. ERROR SCALING WITH OVERROTATION}

To obtain the data of Fig. 4c of the main text, we first calibrate the \( \hat{n} \) and \( \hat{z} \) axes following the calibration procedure described in Sec. II. The voltage-to-rotation angle mapping, \( \theta(V) \), is then used to select rotations which differ from the nominal value by a percentage \( \Delta \), such that \( \theta_i \to (1 \pm \Delta)\theta_i \). The sign of the overrotation is randomly chosen prior to running the blind randomized benchmarking routine. In Fig. 4c, multiple data points for a given value of \( \Delta \) represent distinct instantiations of the randomly-selected sign of the overrotation.

From the definition of gate fidelity for unitary operators [14],

\[
F_g (U, \tilde{U}) = \frac{|\text{Tr} \left( \tilde{U}^\dagger U \right) |^2 / d + 1}{d + 1}, \tag{32}
\]

we can calculate the effect of an implemented unitary, \( \tilde{U} \), relative to the ideal, \( U \), in a system of dimension \( d \). For a given percent angular overrotation, \( \Delta \), we have

\[
U = \exp \left( -i \theta \sigma_n / 2 \right), \tag{33}
\]

\[
\tilde{U} = \exp \left[ -i (\theta + \Delta \theta) \sigma_n / 2 \right]. \tag{34}
\]

Inserting these into the trace of Eqn. 32 gives

\[
|\text{Tr} \left( \tilde{U}^\dagger U \right) |^2 = |\text{Tr} \left[ \exp \left( i \Delta \theta \sigma_n / 2 \right) \right] |^2 = |\text{Tr} \left[ \exp \left( i \Delta \theta / 2 \right) \sigma_n \right] |^2 = 4 \cos^2(\Delta \theta / 2). \tag{35}
\]

Taking \( d = 2 \) and performing a linear expansion of the infidelity, \( 1 - F_g (U, \tilde{U}) \), we obtain the expected scaling of a single gate error from overrotation

\[
\varepsilon_g = 1 - F_g (U, \tilde{U}) = \frac{2}{3} \sin^2(\Delta \theta / 2) \approx \frac{(\Delta \theta)^2}{6} + \mathcal{O}(\Delta \theta)^4. \tag{36}
\]

For an average rotation angle of \( \theta = \pi \) and 2.7 pulses per Clifford, this yields an expected scaling of \( 2.7 \times \frac{\pi^2}{6} \Delta^2 \approx 4.4 \Delta^2 \). In the main text, Fig. 4c, we show that the total error from Blind RB scales as \( 0.0042 + 4.3\Delta^2 \), where the constant offset reflects the intrinsic qubit error in the absence of intentional overrotation.
V. SIMULATIONS OF RANDOMIZED BENCHMARKING EXPERIMENTS

The simulations in Fig. 4d of the main text mimic the experimental procedure quite closely, with accurate models for magnetic and charge noise. The primary differences are that the simulation uses perfect initialization and measurement, and exchange control is implemented as square pulses with noise on exchange rates $J_z(t)$ and $J_n(t)$ but without the pulse distortion from bandwidth limitations that we believe to be present. The contact-hyperfine interactions between electron spins and nuclei ($^{29}$Si and $^{73}$Ge) are approximated as a fluctuating Overhauser field for each electron spin. The simulation uses the full 8-dimensional basis of 3 spins, allowing arbitrary pairwise exchange and local magnetic field vectors on each spin. The simulated experiment is as follows:

1. Preparation – a singlet on two spins is prepared with a third spin with random orientation.

2. Evolve – a blind RB pulse sequence proceeds using the same Clifford constructions as the experiment, with each exchange rotation implemented as a square pulse with duration of 10 ns followed an by idle ($J = 0$) duration $t_{idle}$ as indicated for the experiment or 20 ns when not specified. The effects of the magnetic environment are simulated during all pulse and idle intervals.

3. Measurement – the two-spin singlet probability is directly extracted from the simulated wavefunction.

The simulated Hamiltonian has the following form:

$$H(t) = J_z(t)\mathbf{S_1} \cdot \mathbf{S_2} + J_n(t)\mathbf{S_2} \cdot \mathbf{S_3} + \sum_{k=1}^{3} b_k(t) \cdot \mathbf{S_k},$$

where each $b_k(t)$ is the simulated Overhauser field for dot $k$. Noise on each $J_z(t)$ and $b_k(t)$ is simulated as described below. The simulator approximates evolution with a first-order Magnus expansion:

$$\mathcal{T} \exp \left( -\frac{i}{\hbar} \int_{t_a}^{t_b} H(t)dt \right) \approx \exp \left( -\frac{i}{\hbar} \int_{t_a}^{t_b} H(t)dt \right).$$

Each simulation interval $[t_a, t_b]$ is a single pulse or idle interval between pulses. After a single simulated experiment, the simulator is advanced forward in time equal to the experimental timescale for repeating the same pulse sequence and averaging the single-shot measurements to estimate the population of $|0\rangle$. In order to capture any long-timescale correlations due to low-frequency noise, the schedule of performing RB with multiple random Clifford sequences and multiple averages closely mimics the schedule in the actual experiment.

Noise is introduced in two forms, the Overhauser field that approximates the impact of contact-hyperfine interactions with nuclei and charge noise affecting the exchange interaction. Noise with $1/f$-like power spectrum is simulated through a variant of the Voss-McCartney algorithm, which is essentially an ensemble of gaussian fluctuators with appropriately chosen switching rates, extended to evolution over arbitrary-length intervals. First, the magnetic noise is an independent $1/f$ ensemble for each of the $b_{x}(t)$, $b_{y}(t)$, $b_{z}(t)$ magnetic fields for each electron. The $1/f$ noise spectral density corresponds to observed magnetic noise in this and previous Si/SiGe quantum dot devices. The spectrum is bounded by providing a low-frequency cut-off to white noise below 1 mHz and a high-frequency roll-off to $1/f^2$ for noise above 100 kHz. These timescales are approximately based on experimentally observed averaging times and bounds to the hyperfine and nuclear dipole-dipole Hamiltonian. Simulation results depend very weakly on the lower- and upper-frequency roll-offs. The total integrated noise is chosen to match the observed degree of dephasing observed in the decay of double-dot singlet coherence in this device (shown in Fig. S2a), which is similar to that discussed in Ref. 10. Note that the decay shape and asymptotic singlet probability in Fig. S2a match expectations for a singlet dephasing under slowly varying, independently fluctuating magnetic fields in double dots at low applied magnetic field [15]; in this case the magnetic field is that of the Earth (~50 µT). The corresponding simulation is shown in Fig. S2b.

The second form of simulated noise is charge noise. Each exchange term in Eqn. (37) is simulated as $J(t) = J_0 [1 + \epsilon(t)]$, where $\epsilon(t)$ is a $1/f$ noise process having a high-frequency roll-off at 10 GHz, much faster than any pulsing time, and a low-frequency cut-off slower than any experimental timescale (typically mHz). For each pulse of angle $\theta$ and duration $t_{\text{pulse}}$, $J_0 = \theta/t_{\text{pulse}}$, i.e. the exchange needed for a square pulse to implement the intended rotation. The simulator sets $J_0 = 0$ during idle periods, and only one exchange term is nonzero during pulsing; however, the simulated noise processes are advanced in time to properly include time correlations. Notably, simulating square pulses and $J = 0$ during idle periods does not account for pulse distortion, which we consider to be an important noise source at short idle time, as discussed in the main text. The amplitude of this $1/f$ spectrum is similar to that
FIG. S2. Measurement of singlet dephasing due to charge noise and fluctuating hyperfine-induced magnetic field gradients at low field. A singlet is prepared on dots 1 and 2, allowed to dephase in the presence of a, negligible exchange, or c, 93 MHz exchange between dots 2 and 3, and then measured. b, Simulation of the experiment in a. d, Simulation of the experiment in c.

discussed in Ref. [5], consistent with Rabi-oscillation measurements in this device as shown in Fig. S2c and providing the simulated decay of coherence shown in Fig. S2d.

[1] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley, “Universal quantum computation with the exchange interaction,” Nature 408, 339 (2000).
[2] A. Messiah, Quantum Mechanics, Vol. 2 (North-Holland, 1962) pp. 573–575.
[3] F. Loventhal, “Uniform finite generation of the rotation group,” Rocky Mountain J. Math. 1, 575 (1971).
[4] T. D. Ladd, “Hyperfine-induced decay in triple quantum dots,” Phys. Rev. B 86, 125408 (2012).
[5] M. D. Reed, B. M. Maune, R. W. Andrews, M. G. Borselli, K. Eng, M. P. Jura, A. A. Kiselev, T. D. Ladd, S. T. Merkel, I. Milosavljevic, E. J. Pritchett, M. T. Rakher, R. S. Ross, A. E. Schmitz, A. Smith, J. A. Wright, M. F. Gyure, and A. T. Hunter, “Reduced sensitivity to charge noise in semiconductor spin qubits via symmetric operation,” Phys. Rev. Lett. 116, 110402 (2016).
[6] J. P. Burg, Maximum Entropy Spectral Analysis, Ph.D. thesis, Stanford University (1975).
[7] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, “Randomized benchmarking of quantum gates,” Phys. Rev. A 77, 012307 (2008).
[8] J. T. Muhonen, A. Laucht, S. Simmons, J. P. Dehollain, R. Kalra, F. E. Hudson, S. Freer, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello, “Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking,” J. Phys. Condens. Matter 27, 154205 (2015).
[9] J. Helson, J. J. Wallman, S. T. Flammia, and S. Wehner, “Multi-qubit randomized benchmarking using few samples,” arXiv:1701.04299v2.
[10] K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, “Isotopically enhanced triple-quantum-dot qubit,” Sci. Adv. 1, e1500214 (2015).
[11] A. M. Jones, E. J. Pritchett, E. H. Chen, T. E. Keating, R. W. Andrews, J. Z. Blumoff, L. A. De Lorenzo, K. Eng, S. D. Ha, A. A. Kiselev, S. M. Meenehan, S. T. Merkel, J. A. Wright, L. F. Edge, R. S. Ross, M. T. Rakher, M. G. Borselli, and A. Hunter, “Spin-blockade spectroscopy of Si/SiGe quantum dots,” arXiv:1809.08320.
[12] B. H. Fong and S. M. Wandzura, “Universal quantum computation and leakage reduction in the 3-qubit decoherence free subsystem,” Quantum Info. Comput. 11, 1003 (2011).

[13] J. J. Wallman, M. Barnhill, and J. Emerson, “Robust characterization of leakage errors,” New J. Phys. 18, 043021 (2016).

[14] M. A. Nielsen, “A simple formula for the average gate fidelity of a quantum dynamical operation,” Phys. Lett. A 303, 249 (2002).

[15] K. Schulten and P. G. Wolynes, “Semiclassical description of electron spin motion in radicals including the effect of electron hopping,” J. Chem. Phys. 68, 3292–3297 (1978).