nature physics

Article

https://doi.org/10.1038/s41567-024-02614-w

Assessment of the errors of high-fidelity two-qubit gates in silicon quantum dots

Received: 27 September 2023

Accepted: 12 July 2024

Published online: 20 August 2024

Check for updates

Tuomo Tanttu[1,2] ✉, Wee Han Lim[1,2], Jonathan Y. Huang[1], Nard Dumoulin Stuyck[1,2], Will Gilbert[1,2], Rocky Y. Su[1], MengKe Feng[1,2], Jesus D. Cifuentes[1,2], Amanda E. Seedhouse[1,2], Stefan K. Seritan[3], Corey I. Ostrove[4], Kenneth M. Rudinger[4], Ross C. C. Leon[1,10], Wister Huang[1,11], Christopher C. Escott[1,2], Kohei M. Itoh[5], Nikolay V. Abrosimov[6], Hans-Joachim Pohl[7], Michael L. W. Thewalt[8], Fay E. Hudson[1,2], Robin Blume-Kohout[4], Stephen D. Bartlett[9], Andrea Morello[1], Arne Laucht[1,2], Chih Hwan Yang[1,2], Andre Saraiva[1,2] & Andrew S. Dzurak[1,2] ✉

Achieving high-fidelity entangling operations between qubits consistently is essential for the performance of multi-qubit systems. Solid-state platforms are particularly exposed to errors arising from materials-induced variability between qubits, which leads to performance inconsistencies. Here we study the errors in a spin qubit processor, tying them to their physical origins. We use this knowledge to demonstrate consistent and repeatable operation with above 99% fidelity of two-qubit gates in the technologically important silicon metal-oxide-semiconductor quantum dot platform. Analysis of the physical errors and fidelities in multiple devices over extended periods allows us to ensure that we capture the variation and the most common error types. Physical error sources include the slow nuclear and electrical noise on single qubits and contextual noise that depends on the applied control sequence. Furthermore, we investigate the impact of qubit design, feedback systems and robust gate design to inform the design of future scalable, high-fidelity control strategies. Our results highlight both the capabilities and challenges for the scaling-up of silicon spin-based qubits into full-scale quantum processors.

The fidelities of qubit operations need to be consistent over time and across different qubits to achieve complex quantum computations, such as those in recent groundbreaking research[1–5]. The physical mechanisms behind the entanglement between qubits are key to the success of a two-qubit gate and have a large impact in the performance of algorithms and error correction schemes[6–9]. Exchange-based entangling gates between silicon spin qubits have only recently matured enough to achieve high-fidelity operation[10–14].

The variation of the operational parameters is particularly important in the case of spin qubits owing to their nanometric physical size and nanosecond-scale operation time. While all solid-state systems are subjected to materials noise and disorder, spin qubits probe these imperfections at nearly the atomic scale. In the case of silicon metal-oxide-semiconductor quantum dot qubits[15], this is especially pronounced since the qubits are pressed against the amorphous Si–SiO_2 interface. The issue of the consistency of high-fidelity operations then becomes central to translating this fabrication scalability into a qubit control scalability.

Our goal is to analyse the statistical characteristics and temporal stability of primitive two-qubit gate operations involving electron spins in Silicon metal-oxide-semiconductor quantum dots, which are based on the Heisenberg exchange interaction. Our entangling gates are based on controlling the exchange between spins of electrons in neighbouring quantum dots by pulsing the height of the tunnel

A full list of affiliations appears at the end of the paper. ✉e-mail: t.tanttu@unsw.edu.au; a.dzurak@unsw.edu.au

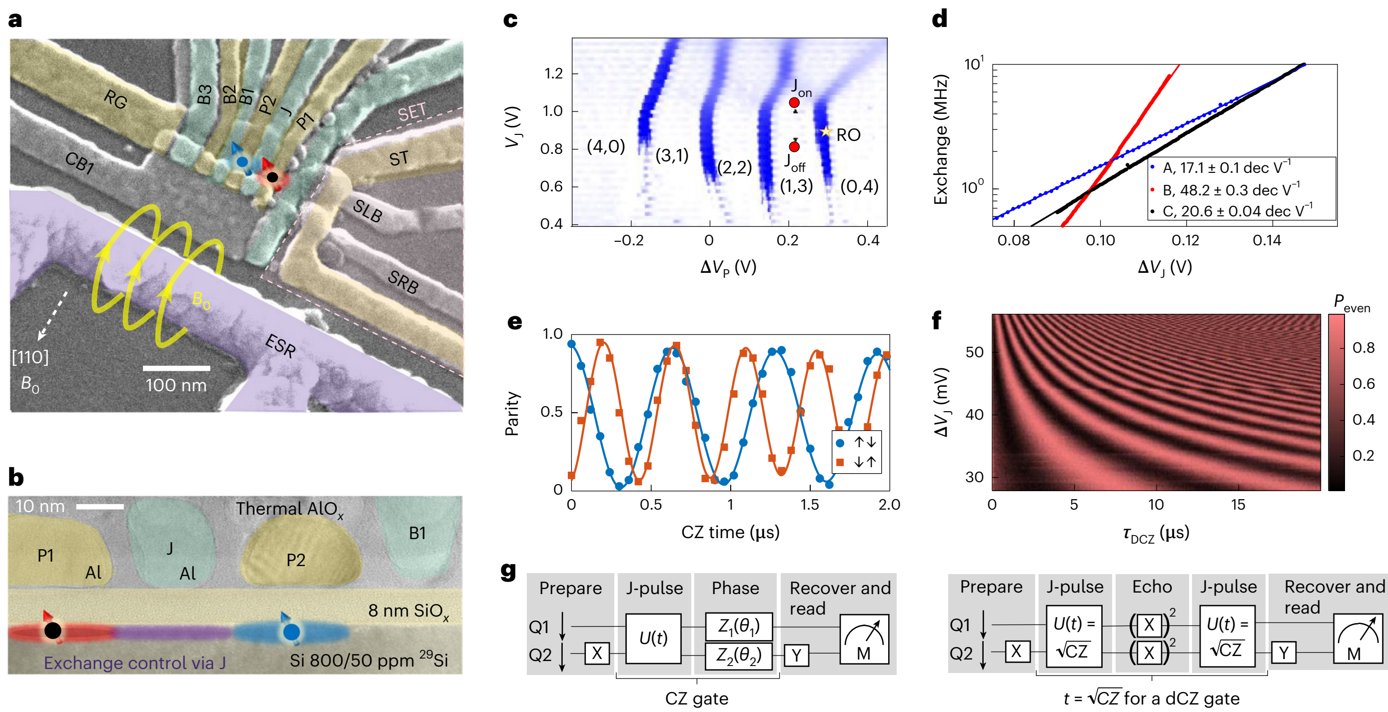


Fig. 1 | Electrostatic quantum dots with tunable exchange. a, False colour scanning electron micrograph (SEM) of a device similar to A and B The qubit dots are accumulated under the plunger (P) gates. The charge state of the double dot is monitored, by the sensor dot which is confined between the barrier gates SLB and SRB. **b**, False colour transmission electron micrograph (TEM) of a cross-section of a device similar to A and B. The shaded area shows the extension of electron wave functions under each dot. **c**, Charge stability map of device A in isolated mode together with the important operation points for qubit operation (J_{on} and J_{off}) and readout (RO). **d**, Exchange energy for devices A, B and C as a function of V_J. The rate of increase of exchange is shown in the legend. **e**, CZ oscillations using the measurement sequence in **g** with two different control qubit initializations at a fixed level of J. **f**, Oscillations of DCZ gate sequence as a function of J-gate pulse time and voltage level. **g**, Pulse sequences used in the experiments for CZ (left) and DCZ (right) in **e** and **f**. Here, 'X' refers to $\frac{\pi}{2}$ rotation around the x axis. For the CZ gate we apply phase θ rotations around z-axis for both qubits to correct for the Stark shift from pulsing the J-gate.

barrier between dots with an interstitial exchange-control electrode. We perform entangling gates with two strategies: a simple square pulse of voltage, which leads to an effective Ising interaction between spins implementing controlled phase (CZ)[16]; or a composite gate consisting of two voltage pulses separated by a microwave pulse that performs single-qubit dynamical decoupling, referred to as decoupled controlled phase (DCZ) gates[17]. We analyse the errors introduced by each of these strategies in detail, leveraging three state-of-the-art methods of validation.

In addition to the multiple validation methods, we verify the consistency of the two-qubit gate operations by reproducing them in three different devices. Two devices (A and B) are nominally identical three-dot chains, having been fabricated in the same batch, with the gate layout shown in Fig. 1a. The third device (C in Extended Data Fig. 1b) has four dots instead, but the same choice of material stack (aluminium gates and thermally grown Al_2O_3). All experiments shown here are based on forming only two of the dots at a time.

The level of isotopic purification of the silicon substrate is also different—800 ppm ^{29}Si for devices A and B and 50 ppm for device C. We do not present results comparing silicon substrate purification levels, but qualitatively the more purified device C was operated with high fidelity without as much need for active feedback on the qubit parameters (only a single parameter, instead of nine or seven in the cases of devices A and B, respectively).

Our electron spin qubits reside under the plunger gates, P1 and P2, and the oscillating B_1 field from a nearby antenna drives the electron spin resonance induced by magnetic field B_0. The exchange interaction is tuned via the voltage on the interstitial exchange gate, J. For measurement, we use a single-electron transistor (SET) that senses charge movement and becomes conditional on the spins when we try to move both electrons into the same dot due to the Pauli exclusion principle. The detailed cross-section image in Fig. 1b displays the active area where the qubits are formed and also reveals the oxide variation and fabrication inconsistencies that lead to variations in device parameters and performance that, in turn, lead to the difficulty in obtaining consistent fidelities across devices[18]. The differences between devices are laid out in Extended Data Table 1.

The isolated mode[19,20] stability map in Fig. 1c for device A with four electrons reveals the charge transitions between dots and allows us to choose the operational points in the device (see Extended Data Fig. 1a,c for devices B and C). We choose a symmetric operation point for our exchange-on voltage to ensure that our two-qubit gate is not sensitive to the detuning noise. The d.c. biases are typically chosen so that exchange-on point is at the high end of our dynamic range (400 mV), allowing us to lower the tunnel rates for the blockade readout.

The spins in all our experiments are read out through the relative parity of the two spins using Pauli spin blockade for spin-to-charge conversion[21,22]. After finding the resonance frequencies of the qubits, we analyse their coherence, Rabi frequencies and noise spectra. In one case, for device A, we use a vector magnet to study the effect of the direction of the constant magnetic field and identify the microscopic origin of the decoherence. Our analysis on device A (see magnetic field dependence measurement in ref. 23) shows that T_2^*s are limited by comparative spin–orbit and ^{29}Si noise. Device B, being from the same batch as device A, is likely to have similar noise limitations. Device C is probably limited mostly by electric noise, thanks to its superior isotopic purity compared with devices A and B. Our T_2^{Hahn} is more strongly limited by the charge rather than hyperfine noise in all devices. To fight the low-frequency components of the intrinsic $1/f$ noise and slow jumps due to hyperfine interactions with ^{29}Si nuclear spins, we apply

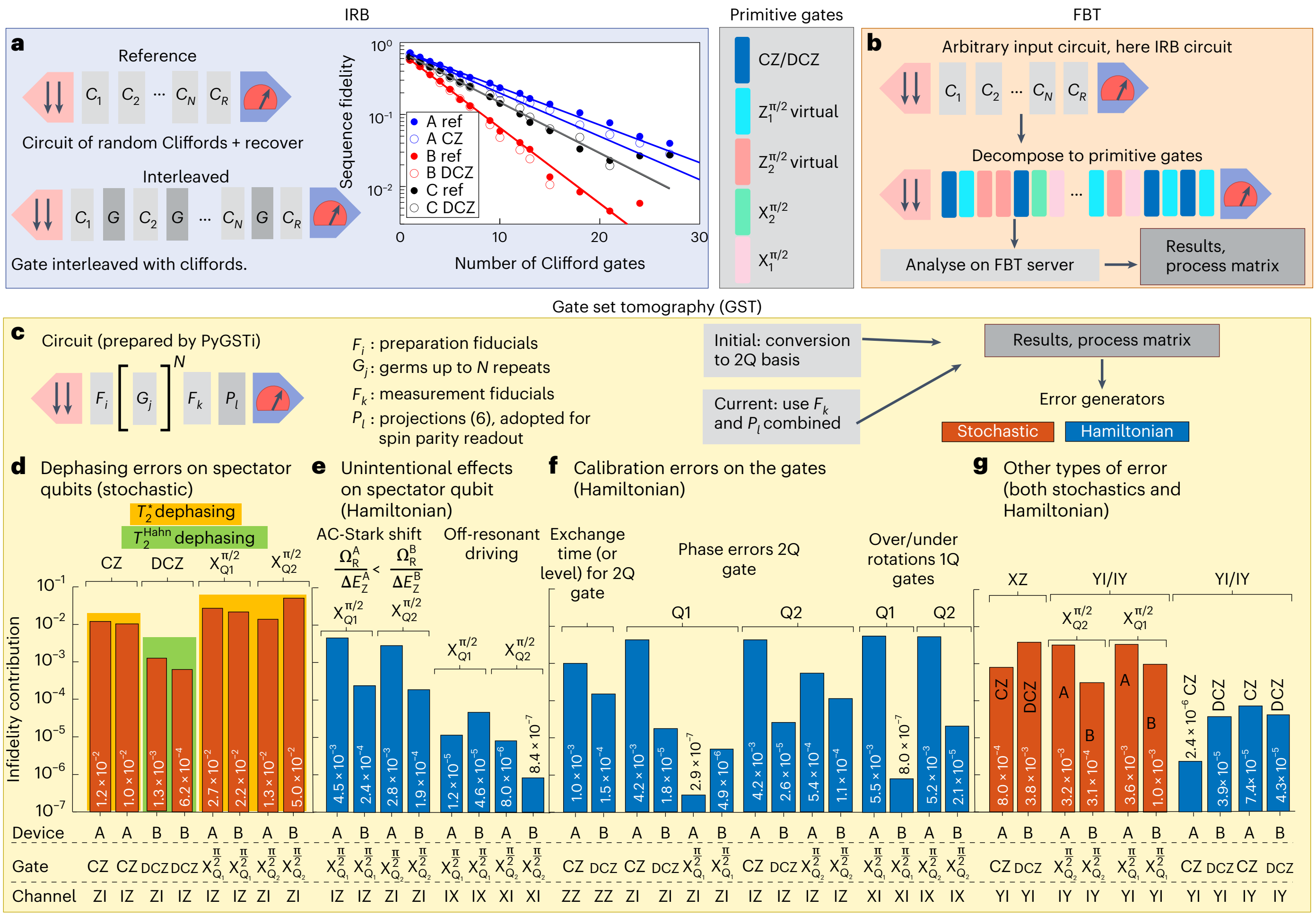


Fig. 2 | Summary of tomographic methods and selection of major identified physical error sources of the gate implementations based on GST for one- and two-qubit gates in devices A and B. a, Measurement sequence principle used in randomized benchmarking. The gate of interest (G) is interleaved with N random Clifford gates (C_i) together with recovery Clifford (C_R) that are composed of five primitive gates each. Recovery probability of the randomized benchmarking sequence as a function of Clifford gates for both interleaved and reference (ref) sequences in all devices. **b**, Simplified FBT workflow from experiment to result. FBT can analyse any gate sequence; IRB is used here as an example. **c**, GST workflow used in our experiment. **d**, Dephasing (stochastic) errors: these noise channels occur due to the T_2^*-like decay during the operation and also non-Markovian contextual error sources (Fig. 3). The reduced error rate between DCZ is due to noise limited by T_2^{Hahn}-like decay instead. **e**, Physical Hamiltonian errors that result from operations such as AC-Stark shift, off-resonant driving or residual exchange during the single qubit operation. **f**, Calibration (systematic) errors, due to the errors in the calibration of the gates. **g**, Other major errors with no major physical attribution.

frequency feedback to devices A and B to keep the microwave control in resonance[24]. This is not necessary for device C.

To run high-fidelity two-qubit gates without being limited by the residual exchange at the nominally off state, we aim at an exchange swing that provides at least a 10^4 ratio between on and off states. This is achievable for all devices with our 400 mV dynamical voltage range of the J gates, as shown in Fig. 1d.

Multi-qubit physical errors

We utilize three different validation methods—interleaved randomized benchmarking (IRB)[25,26], fast Bayesian tomography (FBT)[27] and gate set tomography (GST)[28,29]—illustrated in Fig. 2a–c summarized in Extended Data Table 2 and described in further detail in 'Comparison with other high-fidelity two spin systems in silicon' section in Supplementary Discussion. The estimated process matrices for all gates both from FBT and GST are shown in Extended Data Figs. 2 and 3, respectively. Our tomographic analyses allow us to identify the underlying quantum process and form conjectures about the physical mechanisms causing the error. We explored not only the entangling gates but also the two-qubit processes generated by single qubit gates, which elucidates effects such as the exposure of a qubit to decoherence when idling while the other qubit is being controlled, the effect of crosstalk and contextual errors.

We run a GST experiment in device A with CZ implementation and in device B with DCZ implementation. We generate the circuits for the analysis with special consideration for the measurement effects of the parity readout, as opposed to the more common readout of individual qubits. Two strategies are adopted to that end, one entailing the addition of projections of single qubit states into the parity and repeating the experiments for these projections, and the other by adapting the analysis tool itself (pyGSTi[29]) to handle the measurement effects directly at the analysis stage. Incorporating the native measurement operation explicitly and generating measurement fiducials accordingly is found to be the most efficient strategy (see more details in 'Gate set tomography with parity readout' section in Supplementary Discussion).

The most immediate form of analysis of the errors detected by GST is the breakdown between Hamiltonian and stochastic parts, which can be done through a mathematical framework without assumptions of the underlying mechanisms[30]. The Hamiltonian part of the error is in general associated with calibration issues, unintended Hamiltonian

 https://doi.org/10.1038/s41567-024-02614-w

terms or forms of contextual error that only become evident once long circuits are performed (such as the consistent heating of the chip due to the long string of microwave pulses). Feedback into the gate control parameters allows us to minimize these errors until we achieve consistently high fidelities.

Removing these peculiarities of the qubit performance from the picture, the infidelity becomes dominated by stochastic effects that we can compare across devices (noting that the infidelity contribution of Hamiltonian errors is naturally weaker, with a quadratic dependence, compared with the linear dependence on stochastic errors). In Fig. 2d–g, we have grouped the major error channels resulting from this optimization loop according to their physical interpretation: the dephasing errors, calibration errors, physical errors and other uncategorized errors, respectively.

One striking difference between the CZ and DCZ implementations is that both the stochastic and Hamiltonian IZ and ZI errors are an order of magnitude smaller for DCZ. By its design, DCZ echoes out phase accumulation owing to quasistatic shifts in the spin Larmor frequencies and, hence, suppresses stochastic IZ and ZI errors. It also suppresses Hamiltonian IZ and ZI errors incurred due to spurious Stark shifts created by the voltage pulse on the exchange gate. This is the main reason our DCZ gates perform better in device B even though the coherence time, T_2^*, is worse for both qubits in this device compared with device A. This is the primary contributor to differences in the fidelities between the two implementations.

The unintentional driving errors in Fig. 2e are created by frequency crosstalk. The difference in qubit frequencies, ΔE_Z, is large compared with the Rabi frequencies, Ω_R, but is still relevant when high-fidelity operation is attempted, generating alternating current (AC) Stark shift (Z error) and off-resonant driving (X error). These errors appear Hamiltonian since they are systematic. Calibration errors in Fig. 2f are also Hamiltonian and are caused by the realistic limits in calibration accuracy of the gates, leading to errors such as under/overrotation of the gates and exchange level errors. We also note that the environment circumstances in which calibration is performed are different from those during the circuits, which leads to contextual errors.

We note that some errors, such as the ones in Fig. 2g, remain unexplained. For example, the strong presence of IY Hamiltonian errors even in the CZ implementation (which does not contain any microwave pulses) and the appearance of a stochastic XZ error in both devices. These errors are consistent enough across devices and measurement setups that they instigate future investigations of, as yet, unknown microscopic physical mechanisms impacting spins in quantum dots. Together, these sources of error add to a considerable amount and are part of the impediment to achieving the next level of gate fidelities. All the error channels extracted from GST are shown in Extended Data Figs. 4 and 5.

Another striking observation that we obtain both from GST and FBT in Extended Data Table 3 is that for all devices the single qubit gates are the lowest in fidelity. A more careful GST analysis, in Supplementary Table 3, reveals that on-target fidelity—that is, the fidelity of the operation on one qubit discarding effects on the other qubit—is very high. The idling qubit, however, suffers strong impacts from both dephasing while idling and crosstalk considered in the total error.

A caveat for all of our analysis methods is that they all intrinsically assume that the gates and their associated errors are Markovian processes. We have identified physical effects that do not agree with this assumption[31]. Figure 3b, for instance, shows that the qubit Larmor frequencies shift substantially depending on how long the microwave has been operating—an effect potentially associated with the excitation of the two-level fluctuators in the oxide[32]. This leads to consistent biases in calibration for the gates applied late in a long circuit compared with the same gate applied early—a form of contextual noise[33] that is shown in the Fig. 3a. Another non-Markovian effect is the high-amplitude, low-frequency components in noise stemming from either nuclear

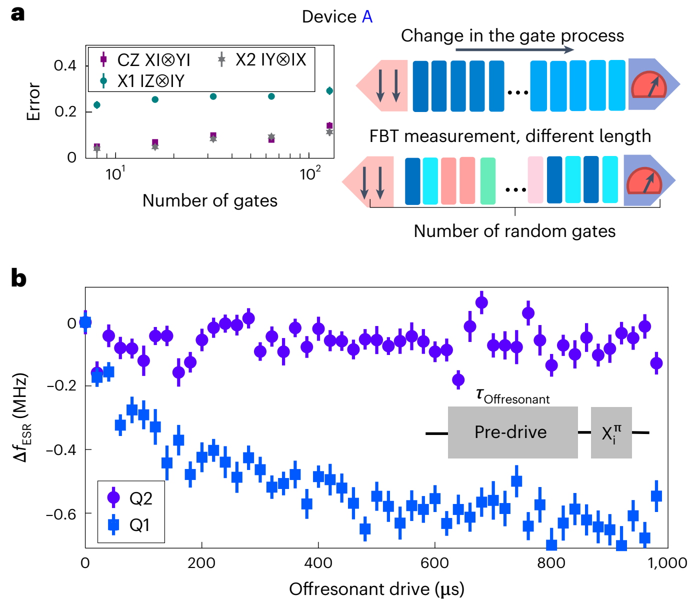


Fig. 3 | Contextual Larmor frequency shift and errors. a, Estimated gate noise channel components XI⊗YI for CZ, IZ⊗IY for $X_1^{\pi/2}$ and IY⊗IX for $X_2^{\pi/2}$ as a function of number of gates in 10,000 randomized sequences with 100 single shots. The shift indicates the changes in the process over the time of one shot of experiment. **b**, The Larmor frequency of both qubits in device A after applying an off-resonant pre-pulse. The Q1 frequency shifts to lower frequencies with longer applied pre-pulse, due to a transient effect possibly similar to those in refs. 32,33,48. This leads to contextual errors observed in **a**. The error bars in both **a** and **b** correspond to the 95% confidence interval.

spins or slow two-level fluctuators that compose the 1/*f* electric noise spectrum[34]. In these circumstances, it becomes evident that differences in the statistical inference approach (frequentist or Bayesian) and the length of circuits (shorter circuits in GST or longer random Cliffords for IRB and FBT) will provide different information about the gates and how they change over time. We also note that, depending on the statistical treatment of the data, these contextual drifts in Hamiltonian error would be interpreted as stochastic noise.

Two-qubit fidelity statistics

Achieving a detailed picture of errors such as the one in Fig. 2 requires a minimum level of stability of high-fidelity operations such that long experiments can be run with consistent results over long periods of time. Our evaluations of two-qubit fidelities from different experimental runs are shown in Fig. 4 and Extended Data Table 3. Most statistics are collected using the simplest validation method: IRB. Based on IRB experiments, we achieved average two-qubit entangling gate fidelities of 98.4% (CZ), 99.37% (DCZ) and 99.76% (DCZ) in devices A, B and C respectively. These numbers indicate sufficient operational fidelity for sustainable error correction[6–9]. More details on the IRB method can be found in ‘Randomized benchmarking’ section in Supplementary Discussion.

We note, however, that non-Markovian effects create notable challenges to the IRB approach and can lead to an overestimated fidelity[35]. In some instances, we found that the circuits performed better when interleaved with DCZ gates, resulting in unphysical estimated fidelities surpassing 100%. On the other hand, relying only on GST experiments exposes our results to instability in the active feedback procedure since they are based on long experiments and a frequentist approach to the statistical treatment of the data. This degradation of the quality of the gates over time is to some extent captured in a Bayesian analysis, as seen by plotting the evolution of the FBT estimates for two-qubit gates during the operation for more than four consecutive hours as seen in the inset of Fig. 4.

https://doi.org/10.1038/s41567-024-02614-w

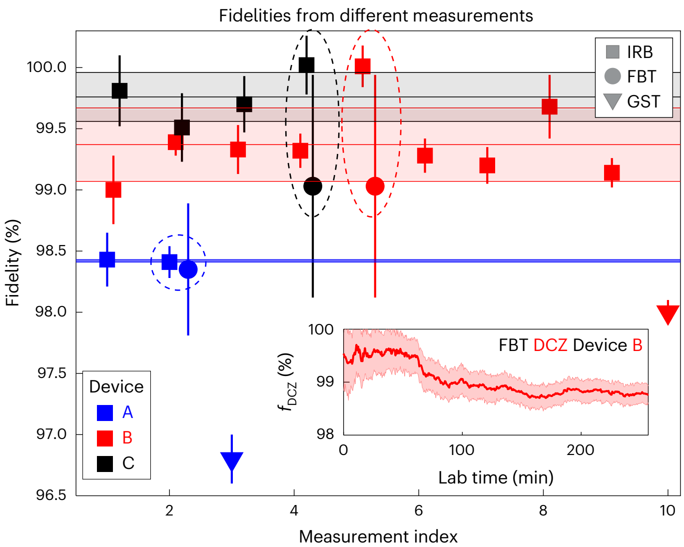


Fig. 4 | Two-qubit gate fidelities. Two-qubit gate fidelities extracted from different error characterization measurements of IRB, and GST. The fidelity of the FBT analysis from certain IRB is the neighbouring data point. This data are also presented in Extended Data Table 3. The error bars are the errors of the fit. The average and the standard deviation of the IRB measurements are indicated with lines and shaded areas around the line. Each IRB run consists of 200 (A) or 500 (B and C) circuit randomizations with 100 single shots for each run. Inset: DCZ fidelity as a function of lab time, where the transient fidelities are extracted from FBT analysis on IRB experiment data. The error bars in the main gfigure and the shaded region in the inset indicate the 95% confidence interval of the fidelity based on the error coming from the fit.

The FBT analysis is performed directly on the outcomes of any arbitrary input circuit (we note that GST may also be adapted for analysing arbitrary circuits). We use the data from IRB runs as input, decomposing all the circuits into their five primitive gates (for additional information, see also 'Implementation of fast Bayesian tomography' section in Supplementary Discussion).

To understand better the experimental runs yielding unphysical fidelities (>100% in B and C), we chose those for the FBT analysis. Tomographic analysis that yields a physical process matrix cannot have above 100% fidelity, and FBT has the additional benefit of allowing the analysis of the process in lab time through Bayesian inference[27]. The fidelities estimated across the full experiment by FBT in Fig. 4 show a very large uncertainty, which can be understood from the plot of the estimated fidelity of the DCZ gate in device B as a function of lab time. Over time, the gate calibration degrades and infidelities more than double.

Implications for large-scale spin quantum computers

Besides the benefits for operation at scale, the consistency of the operations demonstrated here greatly improves the quality of the physical conclusions that can be drawn from experiments. We were able to identify commonalities between gates implemented in different devices and setups and adopting different strategies.

The stability of the operations over time, combined with the theoretical and methodological improvements in process tomography, opens a window into the physics of spin qubit errors. With the possibility of consistently performing high-fidelity one- and two-qubit gates for several hours, we can form hypotheses and test them in a repeatable manner over many hours or days of experiments. Our work not only offers compelling evidence for the microscopic nature of some of the error sources, but it also reveals some unknown processes with, as yet, unclear physical origins.

One of the most important conclusions is that the entangling gate fidelities achieved here can be systematically improved with a combination of better materials to reduce noise[36,37], active use of these tomographic results to recalibrate gates against Hamiltonian errors[12,31] and pulse engineering to reduce the stochastic errors with robust gates and dynamical decoupling[38–41] (see also 'Comparison with other high-fidelity two spin systems in silicon' section in Supplementary Discussion). The fidelity estimates quoted here are far from being fundamentally limited by the physics of spin qubits or by the minimum noise levels in materials.

The scalability of the traditional Loss and DiVincenzo approach[42] to spin qubits can be obscured by the overhead imposed by the strategies required for high-fidelity operation. Due to the relative phases, the number of parameters requiring feedback grows prohibitively fast with the number of qubits if no mitigation strategy is adopted[43]. Moreover, circumventing errors with increasingly convoluted engineered pulses tailored to the idiosyncrasies of each qubit leads to control signals that are hard to generate at scale in an automatized manner. Finally, the degraded performance of idling qubits leads to a steep price in multi-qubit operation. Innovative control solutions might be required to evade these difficulties, including continuously dynamically decoupled driven qubit implementations and pulse shapes aimed at maximizing performance regardless of the particular properties of each qubit[44–47].

Spin qubits in MOS-based quantum dots now join the select group of qubit technologies with two-qubit gate fidelities exceeding the 99% barrier, with the measured average IRB fidelity being 99.17% for three devices with a standard deviation of 0.56%. The prospects for fault-tolerant operation with this platform are further enhanced when the structure of the errors is taken into account; the strong bias towards dephasing errors instead of depolarizing errors opens up the possibility for notable gains in error correction code performance. This will help push the average fidelity up and standard deviation down.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-024-02614-w.

References

1. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
2. Acharya, R. et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature* **614**, 676–681 (2023).
3. Zhong, H.-S. et al. Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
4. Postler, L. et al. Demonstration of fault-tolerant universal quantum gate operations. *Nature* **605**, 675–680 (2022).
5. Madsen, L. S. et al. Quantum computational advantage with a programmable photonic processor. *Nature* **606**, 75–81 (2022).
6. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
7. Stace, T. M., Barrett, S. D. & Doherty, A. C. Thresholds for topological codes in the presence of loss. *Phys. Rev. Lett.* **102**, 200501 (2009).
8. Stace, T. M. & Barrett, S. D. Error correction and degeneracy in surface codes suffering loss. *Phys. Rev. A* **81**, 022317 (2010).
9. Auger, J. M., Anwar, H., Gimeno-Segovia, M., Stace, T. M. & Browne, D. E. Fault-tolerance thresholds for the surface code with fabrication errors. *Phys. Rev. A* **96**, 042316 (2017).
10. Mądzik, M. T. et al. Precision tomography of a three-qubit donor quantum processor in silicon. *Nature* **601**, 348–353 (2022).

11. Noiri, A. et al. Fast universal quantum gate above the fault-tolerance threshold in silicon. *Nature* **601**, 338–342 (2022).
12. Xue, X. et al. Quantum logic with spin qubits crossing the surface code threshold. *Nature* **601**, 343–347 (2022).
13. Mills, A. R. et al. Two-qubit silicon quantum processor with operation fidelity exceeding 99%. *Sci. Adv.* **8**, eabn5130 (2022).
14. Weinstein, A. J. et al. Universal logic with encoded spin qubits in silicon. *Nature* **615**, 817–822 (2023).
15. Veldhorst, M. et al. An addressable quantum dot qubit with fault-tolerant control-fidelity. *Nat. Nanotechnol.* **9**, 981–985 (2014).
16. Veldhorst, M. et al. A two-qubit logic gate in silicon. *Nature* **526**, 410–414 (2015).
17. Watson, T. F. et al. A programmable two-qubit quantum processor in silicon. *Nature* **555**, 633–637 (2018).
18. Cifuentes, J. D. et al. Bounds to electron spin qubit variability for scalable CMOS architectures. *Nat. Commun.* **15**, 4299 (2024).
19. Eenink, H. G. J. et al. Tunable coupling and isolation of single electrons in silicon metal-oxide-semiconductor quantum dots. *Nano Lett.* **19**, 8653–8657 (2019).
20. Yang, C. H. et al. Operation of a silicon quantum processor unit cell above one kelvin. *Nature* **580**, 350–354 (2020).
21. Johnson, A. C., Petta, J. R., Marcus, C. M., Hanson, M. P. & Gossard, A. C. Singlet–triplet spin blockade and charge sensing in a few-electron double quantum dot. *Phys. Rev. B* **72**, 165308 (2005).
22. Seedhouse, A. E. et al. Pauli blockade in silicon quantum dots with spin–orbit control. *PRX Quantum* **2**, 010303 (2021).
23. Cifuentes, J. D. et al. Impact of electrostatic crosstalk on spin qubits in dense CMOS quantum dot arrays. *Phys. Rev. B* (in the press).
24. Stuyck, N. D. et al. Silicon spin qubit noise characterization using real-time feedback protocols and wavelet analysis. *Appl. Phys. Lett.* **124**, 114003 (2024).
25. Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
26. Magesan, E. et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.* **109**, 080505 (2012).
27. Evans, T. et al. Fast bayesian tomography of a two-qubit gate set in silicon. *Phys. Rev. Appl.* **17**, 024068 (2022).
28. Greenbaum, D. Introduction to quantum gate set tomography. Preprint at https://arxiv.org/abs/1509.02921v1 (2015).
29. Nielsen, E. et al. Gate set tomography. *Quantum* **5**, 557 (2021).
30. Blume-Kohout, R. et al. A taxonomy of small markovian errors. *PRX Quantum* **3**, 020335 (2022).
31. Su, R. Y. et al. Characterizing non-Markovian quantum process by fast Bayesian tomography. Preprint at https://arxiv.org/abs/2307.12452 (2023).
32. Undseth, B. et al. Hotter is easier: unexpected temperature dependence of spin qubit frequencies. *Phys. Rev. X* **13**, 041015 (2023).
33. Takeda, K. et al. Optimized electrical control of a Si/SiGe spin qubit in the presence of an induced frequency shift. *npj Quantum Inf.* **4**, 54 (2018).
34. Chan, K. W. et al. Assessment of a silicon quantum dot spin qubit environment via noise spectroscopy. *Phys. Rev. Appl.* **10**, 044017 (2018).
35. Fogarty, M. A. et al. Nonexponential fidelity decay in randomized benchmarking with low-frequency noise. *Phys. Rev. A* **92**, 022326 (2015).
36. Elsayed, A. et al. Low charge noise quantum dots with industrial cmos manufacturing. *npj Quantum Inf.* **10**, 70 (2024).
37. Wan, D. et al. in *Advanced Etch Technology and Process Integration for Nanopatterning XII* Vol. PC12499, PC1249908 (eds Mohanty, N. & Altamirano-Sánchez, E.) (International Society for Optics and Photonics, 2023).
38. Güngördü, U. & Kestner, J. P. Analytically parametrized solutions for robust quantum control using smooth pulses. *Phys. Rev. A* **100**, 062310 (2019).
39. Güngördü, U. & Kestner, J. P. Robust quantum gates using smooth pulses and physics-informed neural networks. *Phys. Rev. Res.* **4**, 023155 (2022).
40. Tang, H. L. et al. Designing globally time-optimal entangling gates using geometric space curves. *Phys. Rev. Appl.* **19**, 044094 (2023).
41. Yang, C. H. et al. Silicon qubit fidelities approaching incoherent noise limits via pulse engineering. *Nat. Electron.* **2**, 151–158 (2019).
42. Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57**, 120–126 (1998).
43. Philips, S. G. J. et al. Universal control of a six-qubit quantum processor in silicon. *Nature* **609**, 919–924 (2022).
44. Vahapoglu, E. et al. Single-electron spin resonance in a nanoelectronic device using a global field. *Sci. Adv.* **7**, eabg9158 (2021).
45. Hansen, I. et al. Pulse engineering of a global field for robust and universal quantum computation. *Phys. Rev. A* **104**, 062415 (2021).
46. Hansen, I. et al. Implementation of an advanced dressing protocol for global qubit control in silicon. *Appl. Phys. Rev.* **9**, 031409 (2022).
47. Seedhouse, A. E. et al. Quantum computation protocol for dressed spins in a global field. *Phys. Rev. B* **104**, 235411 (2021).
48. Freer, S. et al. A single-atom quantum memory in silicon. *Quantum Sci. Technol.* **2**, 015009 (2017).

[1]School of Electrical Engineering and Telecommunications, UNSW, Sydney, New South Wales, Australia. [2]Diraq, Sydney, New South Wales, Australia. [3]Quantum Performance Laboratory, Sandia National Laboratories, Livermore, CA, USA. [4]Quantum Performance Laboratory, Sandia National Laboratories, Albuquerque, NM, USA. [5]School of Fundamental Science and Technology, Keio University, Yokohama, Japan. [6]Leibniz-Institut für Kristallzüchtung, Berlin, Germany. [7]VITCON Projectconsult GmbH, Jena, Germany. [8]Department of Physics, Simon Fraser University, Vancouver, British Columbia, Canada. [9]Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney, New South Wales, Australia. [10]Present address: Quantum Motion Technologies Ltd., London, UK. [11]Present address: ETH Zurich, Zurich, Switzerland. ✉e-mail: t.tanttu@unsw.edu.au; a.dzurak@unsw.edu.au

Methods

Experimental devices

The three devices studied in this work were fabricated using multi-level aluminium gate-stack silicon MOS technology[49,50] on isotopically enriched silicon-28 substrates of 800 ppm residual ^{29}Si (devices A and B) and of 50 ppm residual ^{29}Si (device C). A layer of SiO_2 of ~8 nm was thermally grown above the silicon substrates. The devices are designed with plunger gate width of 30 nm and gate pitch as small as 50 nm. This allows a 20 nm gap between the plunger gates for the J gate.

Measurement setup

Device A was measured in an Oxford Kelvinox 400HA dilution refrigerator. The d.c. bias voltages were generated from Stanford Research Systems SIM928 Isolated Voltage Sources. Gate pulse waveforms were generated by a Quantum Machines (QM) Operator-X+ (OPX+) and combined with d.c. biases using custom linear bias combiners at room temperature.

Devices B and C were measured in a Bluefors XLD400 dilution refrigerator. The d.c. bias voltages were generated with Basel Precision Instruments SP927 DACs. Gate pulse waveforms were generated by a QM OPX and combined with d.c. biases using custom linear bias combiners at the 4K stage.

The SET current of device A was amplified using a room-temperature I–V converter (Basel SP983c) and sampled by a QM OPX. The SET of devices B and C were connected to a tank circuit for reflectometry measurement, with the tone generated by the QM OPX. The return signal was amplified by a Cosmic Microwave Technology CITFL1 LNA at the 4K stage, and a Mini-circuits ZX60-P33ULN+ and a Mini-circuits ZFL-1000LN+ at room temperature, before being digitized and demodulated by the QM OPX.

For all devices, microwave pulses were generated with a Keysight PSG8267D Vector Signal Generator, with in-phase and quadrature (I/Q) and pulse modulation waveforms generated by the QM OPXs.

Data availability

The datasets generated and/or analysed during this study are available in the Zenodo repository at https://doi.org/10.5281/zenodo.12537838 (ref. 51). Source data are provided with this paper.

Code availability

Error taxonomy with pyGSTi was performed with code from ref. 29. The FBT algorithm is given in ref. 27. Analysis scripts are available in the same Zenodo repository as the data. All other supporting algorithms are provided in the paper in the form of equations and diagrams.

References

49. Angus, S. J., Ferguson, A. J., Dzurak, A. S. & Clark, R. G. A silicon radio-frequency single electron transistor. *Appl. Phys. Lett.* **92**, 112103 (2008).
50. Lim, W. H. et al. Observation of the single-electron regime in a highly tunable silicon quantum dot. *Appl. Phys. Lett.* **95**, 242102 (2009).
51. Tanttu, T. Assessment of error variation in high-fidelity two-qubit gates in silicon. *Zenodo* https://doi.org/10.5281/zenodo.12537838 (2024).

Acknowledgements

We thank A. Dickie and S. Serrano for technical help and T. Evans for help with FBT, H. Stemp with GST and A. Noiri with IRB. UNSW team acknowledges support from the Australian Research Council (FL190100167 and CE170100012), the US Army Research Office (W911NF-17-1-0198 and W911NF-23-10092) and the NSW Node of the Australian National Fabrication Facility. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. J.Y.H., W.G., R.S., M.K.F., A.E.S. and J.D.C. acknowledge support from Sydney Quantum Academy. C.O. S.S. K.M.R. and R.B.-K acknowledge support by the US DOE/SC/ASCR through the Quantum Testbed Pathfinder Program. Sandia is a multiprogram laboratory managed and operated by NTESS, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the US DOE's NNSA under contract no. DE-NA-0003525.

Author contributions

W.H.L. and F.E.H. fabricated the devices, with A.S.D.'s supervision, on isotopically enriched ^{28}Si wafers supplied by K.M.I. (800 ppm), N.V.A., H.-J.P. and M.L.W.T. (50 ppm). T.T. and W.H.L. measured initial devices and developed the final J-gate design. T.T., W.H.L. J.Y.H., N.D.S., W.G. and W.H.L. did the experiments, coding and initial analysis, with A.M., A.L., A.S., C.H.Y. and A.S.D.'s supervision. R.C.C.L. and W.H. helped with the initial experimental designs and ideas. R.Y.S. did the FBT analysis with supervision from S.D.B. and A.S. M.K.F. together with C.O., S.S., K.M.R. and R.B.-K. did the GST analysis. A.E.S. and N.D.S. did the feedback analysis. J.D.C. simulated the device structure to guide the gate design with supervision from C.C.E. and A.S. T.T. wrote the manuscript, with input from all authors.

Funding

Open access funding provided through UNSW Library.

Competing interests

A.S.D. is CEO and a director of Diraq Pty Ltd. T.T. W.H.L., N.D.S., W.G., M.F., J.D.C., A.E.S., C.C.E., F.E.H., A.L., C.H.Y., A.S. and A.S.D. declare equity interest in Diraq. The other authors declare no competing interests.

Additional information

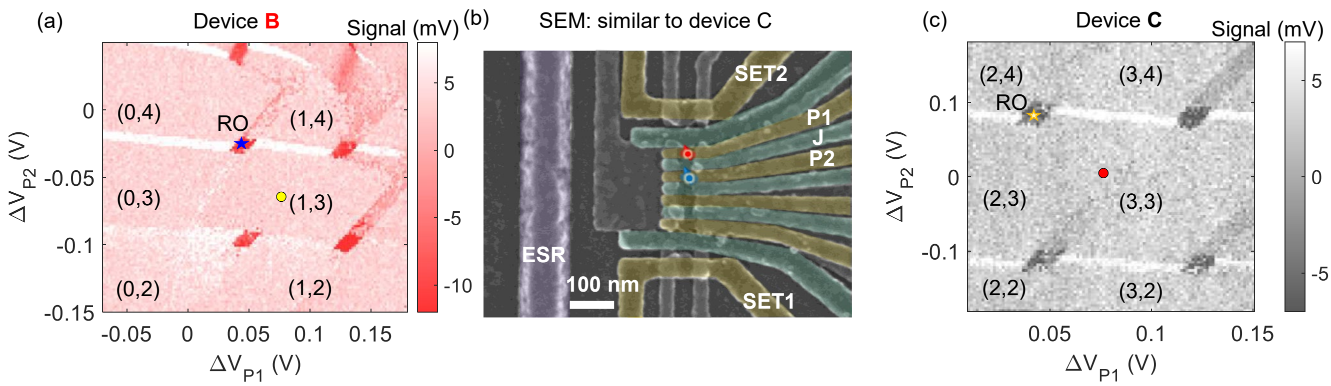
Extended data is available for this paper at https://doi.org/10.1038/s41567-024-02614-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-024-02614-w.

Correspondence and requests for materials should be addressed to Tuomo Tanttu or Andrew S. Dzurak.

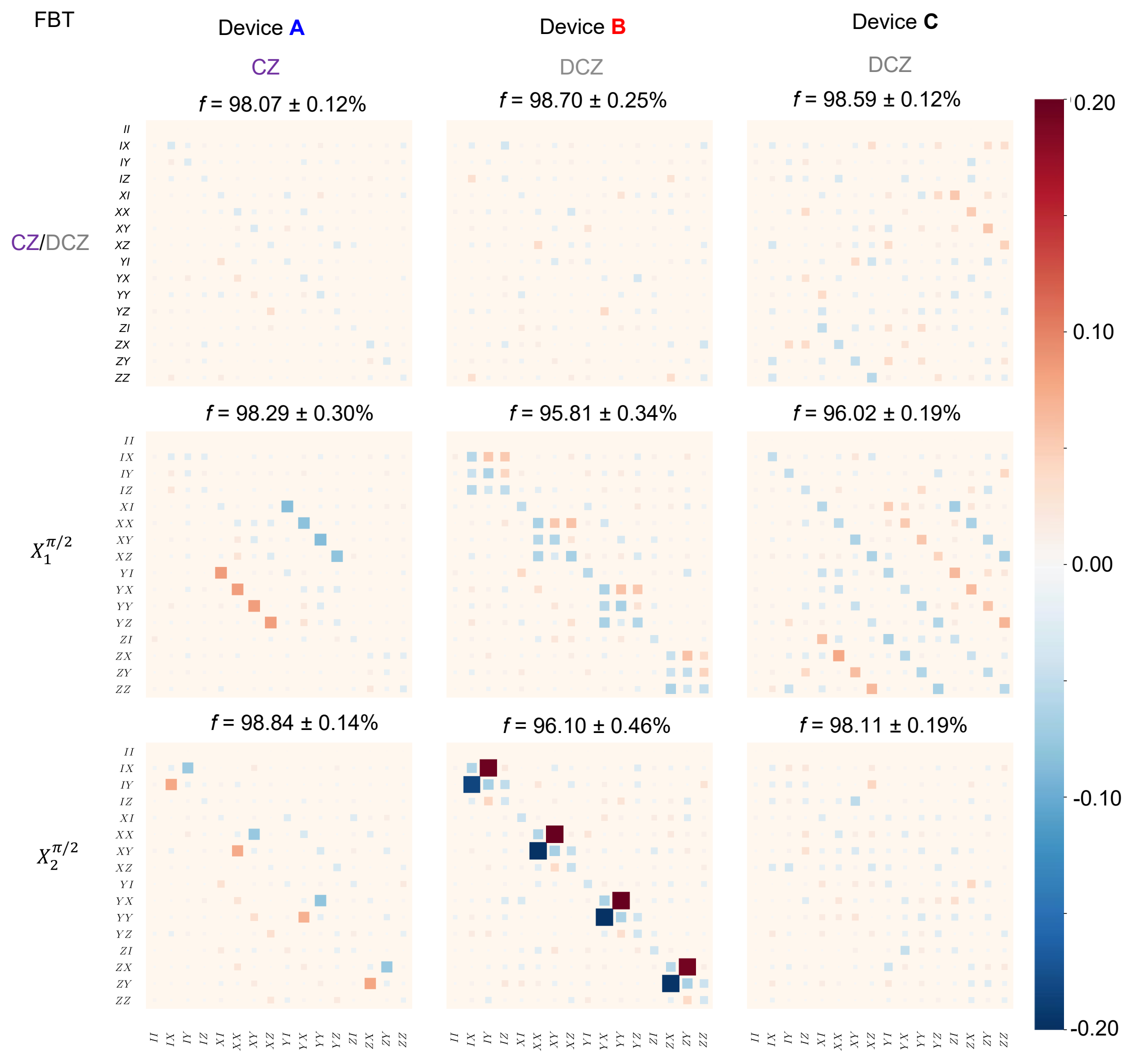
Peer review information *Nature Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41567-024-02614-w

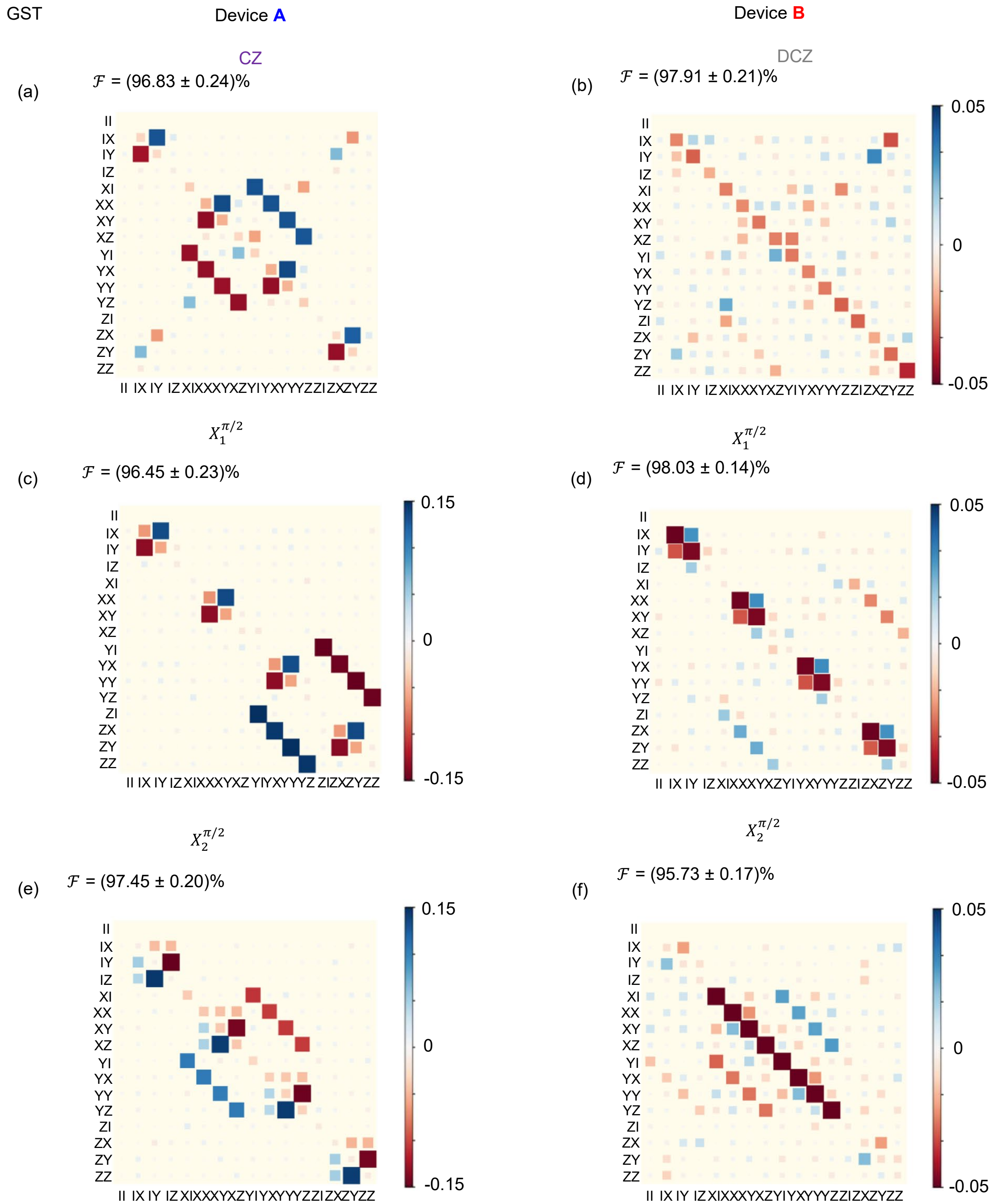


Extended Data Fig. 1 | Device B and C stability maps and SEM of a device similar to C. a, Device B stability map. Operation point (coloured circle) and readout point (RO) are indicated, together with the charge occupancies. **b**, Scanning electron micrograph (SEM) of a device similar to C. **c**, Device C stability map. Indication same as in a.

https://doi.org/10.1038/s41567-024-02614-w

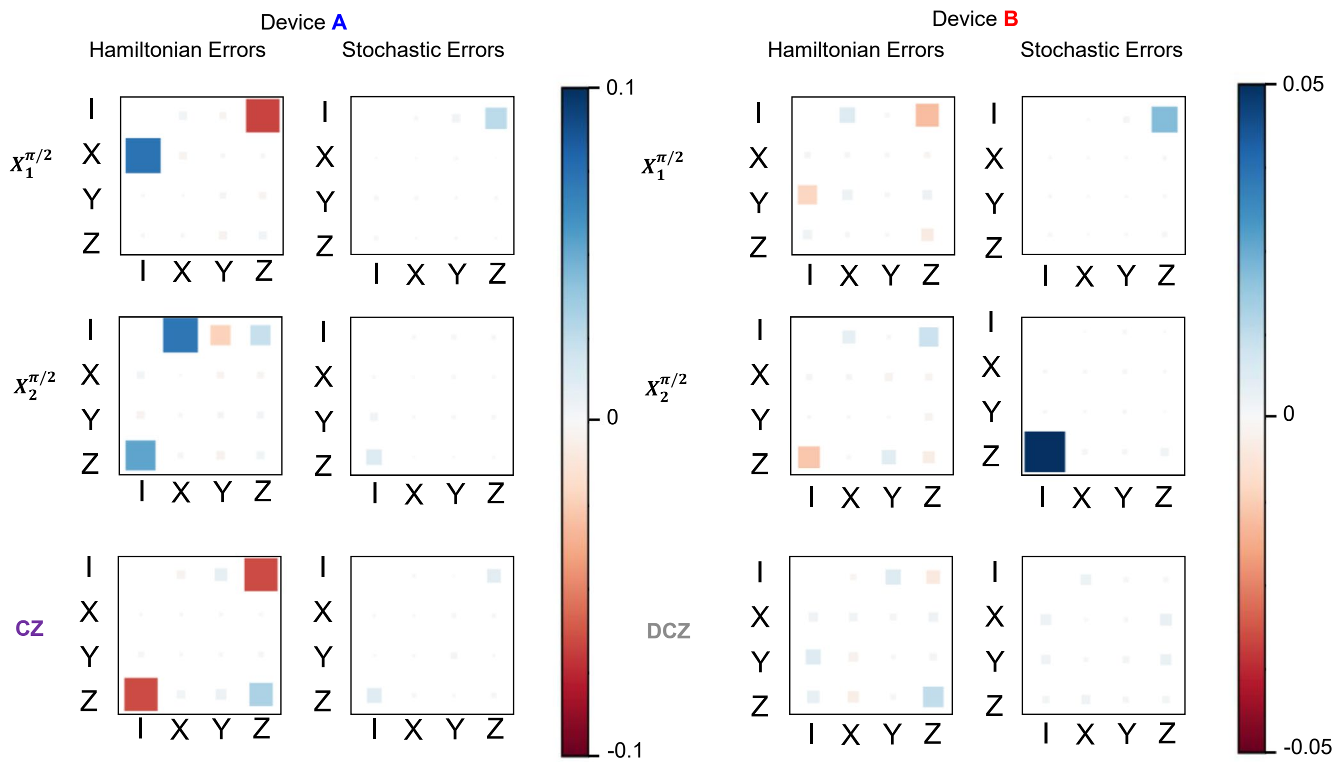


Extended Data Fig. 2 | Error generators from FBT analysis for all three devices, for single and two qubit gates. FBT estimated gate error generators for each device are analyzed from IRB experiments individually. Plotted error generators are the posterior of the last update of each analysis. Fidelity and its uncertainty (with 95% confidence interval) are calculated by re-sampling with the Gaussian statistics of the final estimated results.



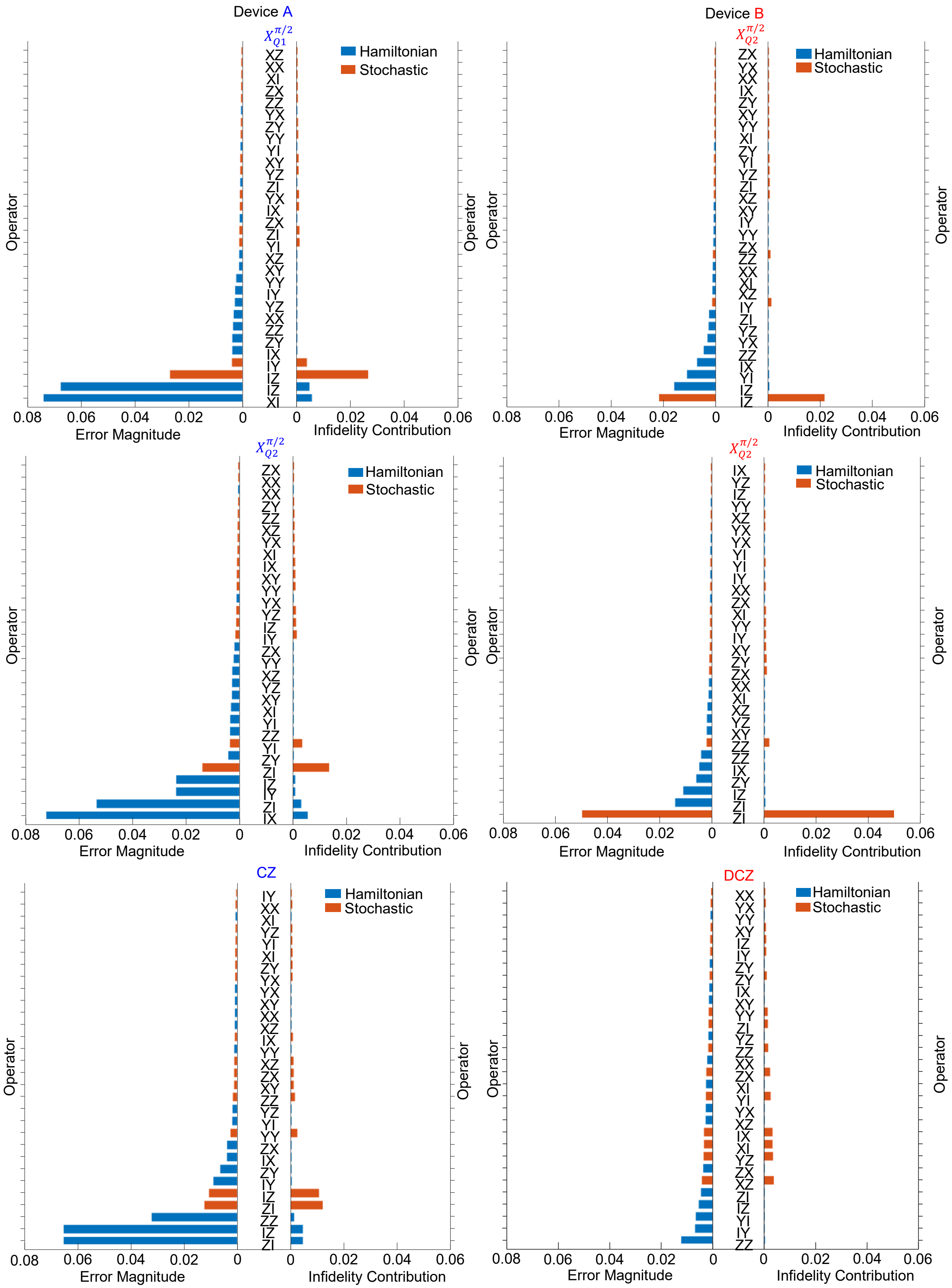
Extended Data Fig. 3 | Hamiltonian and Stochastic error generators from GST reports for devices A and B. **a**, Full error generator for device A, CZ gate. **b**, Full error generator for device B, DCZ gate. **c**, Full error generator for device A, $X_1^{\pi/2}$. **d**, Full error generator for device B, $X_1^{\pi/2}$. **e**, Full error generator for device A, $X_2^{\pi/2}$. **f**, Full error generator for device B, $X_2^{\pi/2}$.

https://doi.org/10.1038/s41567-024-02614-w



Extended Data Fig. 4 | Summary of all the error channels GST components. a, Colour bar ranges from -0.1 to +0.1. **b**, Colour bar ranges from -0.05 to +0.05.

https://doi.org/10.1038/s41567-024-02614-w



Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41567-024-02614-w

Extended Data Fig. 5 | Gate Set Tomography for devices A and B. Error channels magnitudes and the infidelity contributions for all the physical gates for devices A and B. The error channels are listed in order of increasing magnitude in the error magnitude. Note that to get the infidelity contribution from error magnitude the Hamiltonian errors are squared, and stochastic errors are the same. A selection of the largest error magnitude components is plotted in the main text in Fig. 2d,e,f,g. These are extracted from the process matrices in Extended Data Fig. 5.

Extended Data Table 1 | Device parameters

Device	Device **A**		Device **B**		Device **C**	
$T_2^*(\mu s)$ Q1,Q2	5.5 ± 0.3	6.5 ± 0.3	3.0 ± 0.3	2.6 ± 0.3	5.5 ± 0.3	5.5 ± 0.3
$T_2^{\mathrm{Hahn}}(\mu s)$ Q1,Q2	51 ± 3	60 ± 3	39 ± 3	46 ± 3	132 ± 3	72 ± 3
$T_2^{\mathrm{Rabi}}(\mu s)$ Q1,Q2	55 ± 5	64 ± 5	26 ± 3	56 ± 3	131 ± 5	100 ± 5
$X^{\pi/2}(\mu s)$ Q1,Q2	0.520	0.400	0.200	0.380	0.326	0.200
1Q RB fidelity	99.5%	99.6%	99.6%	99.7%	-	-
ΔE_Z (MHz)	10.8		22.7		12.95	
J-gate turn on (dec/V)	17.13		45.15		20.64	
J-gate swing (mV)	330.4		72.24		52.48	
Exchange time CZ/DCZ (ns)	256		500		990	
Exchange on (kHz)	1950		1000		505	
Exchange off (Hz)	4.2		550		$<$13700	
Initialization	T^- (Heralded)		↑↓		T^- (Heralded)	
B_0 and orientation	0.8274 T, [100]		0.893 T, [110]		0.790 T, [110]	
Charge configuration	(1,3) Isolated		(1,3) Reservoir		(3,3) Reservoir	
Parity readout visibility	98 %		99 %		95 %	
GST gauge *invariant* SPAM fidelity	97.6±0.9 %		98.4±0.4 %		-	
Single shot time (μs)	1000		100		200	
Residual ^{29}Si	800 ppm		800 ppm		50 ppm	
Design - number of dots	three		three		four	
Feedbacks	$I_{\mathrm{SET}}, \epsilon, f_{\mathrm{ESR}}^{\mathrm{Qi}}, P_{Xi}^{\mathrm{Qi}}, V_{\mathrm{CZ}}, \varphi_{\mathrm{CZ}}^{\mathrm{Qi}}$		$I_{\mathrm{SET}}, \epsilon, f_{\mathrm{ESR}}^{\mathrm{Qi}}, P_{Xi}^{\mathrm{Qi}}, V_{\mathrm{DCZ}}$		I_{SET}	

Comparison table of the characteristics and operation principles of all the devices used in this experimental campaign.

https://doi.org/10.1038/s41567-024-02614-w

Extended Data Table 2 | Error characterization methods used in this paper

Method	Measurement	Analysis	process matrix
IRB	Randomised Cliffords	decay fit	no
FBT	Any gate sequence	Bayesian estimation	yes
GST	Periodic info-complete fiducials	Maximum-likelihood	yes

Summary of the methods used in this paper.

Extended Data Table 3 | Fidelities extracted from several experimental error characterization runs

Run/Gate	Clifford (%)	CZ/DCZ(%)	$X_1^{\pi/2}$ (%)	$X_2^{\pi/2}$ (%)
IRB_1 **A**	90.90 ± 0.10	98.47 ± 0.22	-	-
IRB_2 **A**	90.85 ± 0.06	98.44 ± 0.13	-	-
mean,std**A**	90.87, 0.04	98.45, 0.01	-	-
IRB_3 **B**	86.95 ± 0.19	99.00 ± 0.28	-	-
IRB_4 **B**	89.65 ± 0.06	99.39 ± 0.11	-	-
IRB_5 **B**	85.06 ± 0.10	99.33 ± 0.20	-	-
IRB_6 **B**	84.96 ± 0.09	100.03 ± 0.17	-	-
IRB_7 **B**	86.13 ± 0.07	99.32 ± 0.14	-	-
IRB_8 **B**	87.52 ± 0.07	99.28 ± 0.14	-	-
IRB_9 **B**	87.29 ± 0.08	99.20 ± 0.15	-	-
IRB_{10} **B**	86.90 ± 0.11	99.68 ± 0.26	-	-
IRB_{11} **B**	87.19 ± 0.07	99.14 ± 0.12	-	-
mean,std **B**	86.85, 1.51	99.37, 0.29	-	-
IRB_{12} **C**	88.18 ± 0.12	99.82 ± 0.29	-	-
IRB_{13} **C**	85.94 ± 0.09	99.51 ± 0.28	-	-
IRB_{14} **C**	90.52 ± 0.11	99.70 ± 0.23	-	-
IRB_{15} **C**	89.04 ± 0.11	100.01 ± 0.24	-	-
mean,std **C**	88.42, 1.92	99.76, 0.21	-	-
FBT **A**	-	98.35 ± 0.54	98.41 ± 0.89	99.09 ± 0.78
FBT **B**	-	99.03 ± 0.91	95.92 ± 0.78	96.58 ± 0.66
FBT **C**	-	99.04 ± 0.54	97.00 ± 1.16	97.72 ± 0.87
GST **A**	-	96.8 ± 0.2	96.5 ± 0.2	97.5 ± 0.2
GST **B**	-	98.0 ± 0.1	97.9 ± 0.2	95.7 ± 0.2

FBT results are from one of the IRB runs indicated by italics. Mean and standard deviation (std) are calculated from all IRB runs for a given device. The error bars for IRB come from the error bars of the fit and for FBT we report the average fidelity over the experiment with error bars being the 2σ variation over the experiment. Clifford fidelity is the average fidelity of a uniformly random 2-qubit Clifford operation. Two-qubit Clifford operations are compiled into native gates, and each one includes, on average, 1.5 CZ/DCZ gates and 1.8 single-qubit $X_{\frac{\pi}{2}}$ gates on each of the two qubits. For GST, the error bars are given by the error of the log-likelihood fit. Some of these are also plotted in the main text Fig. 4.