Precision tomography of a three-qubit electron-nuclear quantum processor in silicon

Mateusz T. Mądzik[*†1,2], Serwan Asaad[*‡1,2], Akram Youssry[3,4], Benjamin Joecker[1,2], Kenneth M. Rudinger[5], Erik Nielsen[5], Kevin C. Young[5], Timothy J. Proctor[5], Andrew D. Baczewski[6], Arne Laucht[1,2], Vivien Schmitt[§1,2], Fay E. Hudson[1], Kohei M. Itoh[7], Alexander M. Jakob[8,2], Brett C. Johnson[8,2], David N. Jamieson[8,2], Andrew S. Dzurak[1], Christopher Ferrie[3], Robin Blume-Kohout[5], and Andrea Morello [¶1,2]

[1]School of Electrical Engineering and Telecommunications, UNSW Sydney, Sydney, NSW 2052, Australia

[2]Centre for Quantum Computation and Communication Technology

[3]Centre for Quantum Software and Information, University of Technology Sydney, Ultimo, NSW 2007, Australia

[4]Department of Electronics and Communication Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt

[5]Quantum Performance Laboratory, Sandia National Laboratories, Albuquerque, NM 87185 and Livermore, CA 94550, USA

[6]Center for Computing Research, Sandia National Laboratories, Albuquerque, NM 87185, USA

[7]School of Fundamental Science and Technology, Keio University, Kohoku-ku, Yokohama, Japan

[8]School of Physics, University of Melbourne, Melbourne, VIC 3010, Australia

Nuclear spins were among the first physical platforms to be considered for quantum information processing[1, 2], because of their exceptional quantum coherence[3] and atomic-scale footprint. However, their full potential for quantum computing has not yet been realized, due to the lack of methods to link nuclear qubits within a scalable device combined with multi-qubit operations with sufficient fidelity to sustain fault-tolerant quantum computation. Here we demonstrate universal quantum logic operations using a pair of ion-implanted ^{31}P nuclei in a silicon nanoelectronic device. A nuclear two-qubit controlled-Z gate is obtained by imparting a geometric phase to a shared electron spin[4], and used to prepare entangled Bell states with fidelities up to 94.2(2.7)%. The quantum operations are precisely characterised using gate set tomography (GST)[5], yielding one-qubit average gate fidelities up to 99.95(2)%, two-qubit average gate fidelity of 99.37(11)% and two-qubit preparation/measurement fidelities of 98.95(4)%. These three metrics indicate that nuclear spins in silicon are approaching the performance demanded in fault-tolerant quantum processors [6]. We then demonstrate entanglement between the two nuclei and the shared electron by producing a Greenberger-Horne-Zeilinger three-qubit state with 92.5(1.0)% fidelity. Since electron spin qubits in semiconductors can be further coupled to other electrons[7, 8, 9] or physically shuttled across different locations[10, 11], these results

*These two authors contributed equally.

†Currently at QuTech, Delft University of Technology, 2628 CJ Delft, The Netherlands.

‡Currently at Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, and Microsoft Quantum Lab Copenhagen, Copenhagen, Denmark.

§Currently at Univ. Grenoble Alpes, Grenoble INP, CEA, IRIG-PHELIQS, F-38000 Grenoble, France.

¶To whom correspondence should be addressed; E-mail: a.morello@unsw.edu.au

establish a viable route for scalable quantum information processing using nuclear spins.

Nuclear spins are the most coherent quantum systems in the solid state [3, 12], owing to their extremely weak coupling to the environment. In the context of quantum information processing, the long coherence is associated with record single-qubit gate fidelities [13]. However, the weak coupling poses a challenge for multi-qubit logic operations. Using spin-carrying defects in diamond [14] and silicon carbide [15], this problem can be addressed by coupling multiple nuclei to a common electron spin, thus creating quantum registers that can sustain small quantum logic operations and error correction [16]. Exciting progress is being made on linking several such defects via optical photons [17, 18].

Still missing, however, is a pathway to exploit the atomic-scale dimension of nuclear spin qubits to engineer scalable quantum processors, where densely-packed qubits are integrated and operated within a semiconductor chip [19]. This requires entangling the nuclear qubits with electrons that can either be physically moved, or entangled with other nearby electrons. It also requires interspersing the electron-nuclear quantum processing units with spin readout devices [20]. Here we show experimentally that silicon - the material underpinning the whole of modern digital information technology - is the natural system in which to develop dense nuclear spin based quantum processors [1].

One electron – two nuclei quantum processor

The experiments are conducted on a system of two ^{31}P donor atoms, introduced in an isotopically purified ^{28}Si substrate by ion implantation (see Methods). A three-qubit processor is formed by using an electron (e) with spin $S = 1/2$ (basis states $|\uparrow\rangle, |\downarrow\rangle$) and two nuclei (Q1, Q2) with spin $I = 1/2$ (basis states $|\Uparrow\rangle, |\Downarrow\rangle$). Metallic structures on the surface of the chip provide electrostatic control of the donors, create a single-electron transistor (SET) charge sensor, and deliver microwave and radiofrequency signals through a broadband antenna (Fig. 1a, Extended Data Fig. 1). With this setup, we can perform single-shot electron spin readout [20], and high fidelity ($\approx$ 99.9%) single-shot quantum nondemolition readout of the nuclear spins [21], as well as nuclear magnetic resonance (NMR) and electron spin resonance (ESR) [22] on all spins involved (see Methods).

The ESR spectra in Fig. 1c exhibit four resonances. This means that the ESR frequency depends upon the state of two nuclei, to which the electron is coupled by contact hyperfine interactions $A_1 \approx 95$ MHz and $A_2 \approx 9$ MHz. We adopt labels where, for instance, $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ represents the frequency at which the electron spin undergoes transitions conditional on the two nuclear spin qubits being in the $|\mathrm{Q}_1\mathrm{Q}_2\rangle = |\Downarrow\Downarrow\rangle$ state, and so on. The values of A_1, A_2 can be independently checked by measuring the frequencies $\nu_{Q1|\downarrow}, \nu_{Q2|\downarrow}$ at which each nucleus responds while the electron is in the $|\downarrow\rangle$ state (Supplementary Information S1). The hyperfine-coupled electron could either be the first or the third electron bound to the donor cluster. Since its spin relaxation time $T_{1\mathrm{e}}$ is three orders of magnitude shorter than expected from a one-electron system (Extended Data Fig. 3), we interpret the ESR spectrum in Fig. 1c as describing the response of the third electron bound to a 2P donor system.

An effective-mass calculation of the wavefunction of the third electron in a 2P system (see Methods) reproduces the observed values of A_1 and A_2 by assuming donors spaced 6.5 nm apart, and subjected to an electric field 2 mV/nm that pulls the electron wavefunction more strongly towards donor 1 (Fig. 1b). The ^{31}P nuclei in this 2P cluster are spaced more widely than those produced by scanning probe lithography [8, 23], where the sub-nanometre inter-donor spacing causes a strongly anisotropic hyperfine coupling, which randomizes the nuclear spin state each time the electron is removed from the cluster for spin readout [24]. Here, instead, the probability of flipping a nuclear spin by electron ionisation is of order 10^{-6} (Extended Data Fig. 5), meaning that our nuclear readout is almost perfectly quantum nondemolition.

Nuclear two-qubit operations

We first consider the two ^{31}P nuclear spins as the qubits of interest. One-qubit logic operations are trivially achieved by NMR pulses [21] (see Methods), where $A_1 \neq A_2$ provides the spectral selectivity to address each qubit individually (Fig. 1c). Two-qubit operations are less trivial, since the nuclei are not directly coupled to each other (Supplementary Information S1 and S9). They are, however, hyperfine-coupled to the same electron. This allows the implementation of a geometric two-qubit controlled-Z (CZ) gate [4, 16].

When a quantum two-level system is made to trace a closed trajectory on its Bloch sphere, its quantum state acquires a geometric phase equal to half the solid angle enclosed by the trajectory [25]. Fig. 1d illustrates how an electron 2π-pulse at the frequency $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ (see Fig. 1d) constitutes a nuclear CZ 2-qubit gate. Starting from the state $|\Downarrow\rangle \otimes (|\Downarrow\rangle + |\Uparrow\rangle)/\sqrt{2} \equiv (|\Downarrow\Downarrow\rangle + |\Downarrow\Uparrow\rangle)/\sqrt{2}$, the electron

Figure_1/Figure_1.pdf

Fig. 1 | Operation of a one-electron – two-nuclei quantum processor. **a,** Artist's impression of a pair of ^{31}P nuclei (red), asymmetrically coupled to the same electron (blue). The spins are controlled by oscillating magnetic fields (yellow) generated on-chip. **b,** Effective-mass calculation of the wavefunction $\psi(y,z)$ of the third electron on the 2P cluster. The observed values of hyperfine coupling are well reproduced by assuming a 6.5 nm spacing between the donors. **c,** Experimental NMR spectrum of the ^{31}P nuclei (top) and ESR spectrum of the shared electron (bottom) at $B_0 = 1.33$ T, along with energy level diagram (right) of the eight-dimensional Hilbert space (spacings not to scale). The spectra yield the hyperfine couplings $A_1 \approx 95$ MHz and $A_2 \approx 9$ MHz between the electron and the nuclear qubits Q1, Q2. **d,** Implementation of a geometric two-qubit CZ gate. A conditional π phase shift is acquired when a 2π rotation is applied on the electron spin at frequency $\nu_{\mathrm{e}|\Downarrow\Downarrow}$, i.e. conditional on the nuclear spins being $|\Downarrow\Downarrow\rangle$. This operation corresponds to the CZ gate on the nuclei when restricted to the electron $|\downarrow\rangle$ subspace.

$X_{2\pi}$ pulse at $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ introduces a phase factor $e^{i\pi} = -1$ to the $|\Downarrow\Downarrow\rangle$ branch of the superposition, resulting in the state $(-|\Downarrow\Downarrow\rangle + |\Downarrow\Uparrow\rangle)/\sqrt{2} \equiv |\Downarrow\rangle \otimes (-|\Downarrow\rangle + |\Uparrow\rangle)/\sqrt{2}$, i.e. a rotation of Q2 by 180 degrees around the z-axis of its Bloch sphere, which is the output of a CZ operation. Conversely, if the initial state of Q1 were $|\Uparrow\rangle$, the pulse at $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ would have no effect on the electron, leaving the nuclear qubits unaffected.

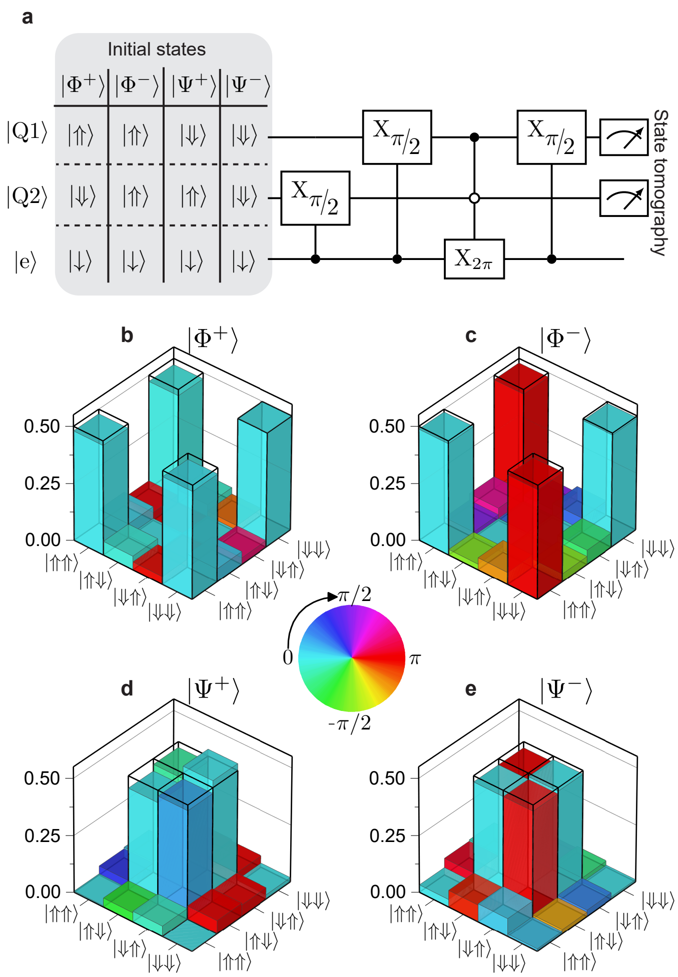
A nuclear controlled-NOT (CNOT) gate is obtained by sandwiching the CZ gate between a nuclear $-\pi/2$ and $\pi/2$ pulse (Extended Data Fig. 6a). Applying an ESR $X_{2\pi}$ pulse at $\nu_{\mathrm{e}|\Uparrow\Downarrow}$ transforms the sequence into a zero-CNOT gate, i.e. a gate that flips Q2 when Q1 is in the $|0\rangle \equiv |\Uparrow\rangle$ state (Extended Data Fig. 6b, and Supplementary Information S2).

We apply this universal gate set (Fig. 2a) to produce each of the four maximally-entangled Bell states of the two nuclear spins, $|\Phi^{\pm}\rangle = (|\Downarrow\Downarrow\rangle \pm |\Uparrow\Uparrow\rangle)/\sqrt{2}$ and $|\Psi^{\pm}\rangle = (|\Downarrow\Uparrow\rangle \pm |\Uparrow\Downarrow\rangle)/\sqrt{2}$. We reconstruct the full density matrices of the Bell states using maximum likelihood quantum state tomography [26] (Supplementary Information S3). The reconstructed states (Fig. 2f) have fidelities of up to 94.2(2.7)%, and concurrences as high as 0.93(4), proving the creation of genuine two-qubit entanglement. Here and elsewhere, error bars indicate 1σ confidence intervals. Bell fidelities and concurrences are calculated without removing state preparation and measurement (SPAM) errors (Extended Data Fig. 10).

Gate set tomography

We used a customized, efficient gate set tomography (GST) [27, 28, 5] analysis (see Methods, and Supplementary Information S4, S5, S8) to investigate the quality of six logic operations on two nuclear-spin qubits: $X_{\pi/2}$ and $Y_{\pi/2}$ rotations on Q1 and Q2, an additional $Y_{-\pi/2}$ rotation on Q2, and the entangling CZ gate. No two single-qubit operations are ever performed in parallel. GST probes these six logic operations and reconstructs a full two-qubit model for their behavior. Earlier experiments on electron spins in silicon used randomized benchmarking (RB) [29, 30] to extract a single number for the average fidelity of all logic operations. Characterising specific gates required "interleaved" RB, which can suffer systematic errors [31, 32]. Most importantly, RB does not reveal the cause or nature of the errors. Our GST method enables measuring each gate's fidelity to high precision, distinguishing the contributions of stochastic and coherent errors, and separating local errors (on the target qubit) from crosstalk errors (on, or coupling to, the undriven spectator qubit).

GST estimates a two-qubit process matrix for each logic operation ($G_i : i = 1 \ldots 6$) using maximum likelihood estimation. We represent each G_i as the composition of its ideal target unitary process ($\mathbb{G}_i$) with an error process written in terms of a Lindbladian generator ($\mathbb{L}_i$): $G_i = e^{\mathbb{L}_i}\mathbb{G}_i$. Each gate's error generator (EG) can be written as a linear combination of independent elementary EGs that describe distinct kinds of error [33]. Each elementary EG's coefficient in $\mathbb{L}_i$ is the rate (per gate) at which that error builds up. Any Markovian error process can be described using just four kinds of elementary EGs: Hamiltonian (H), indexed by a single two-qubit Pauli operator, cause coherent or unitary errors (e.g., H_{ZZ} gen-



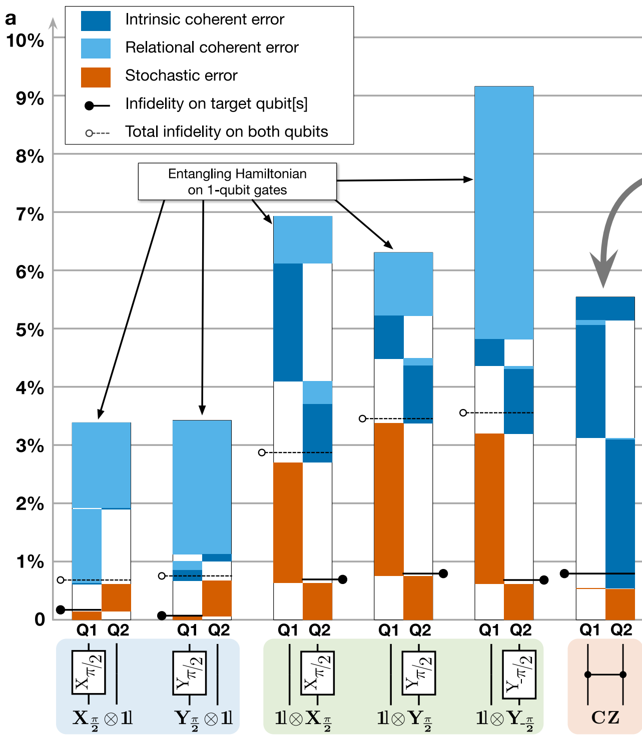
f

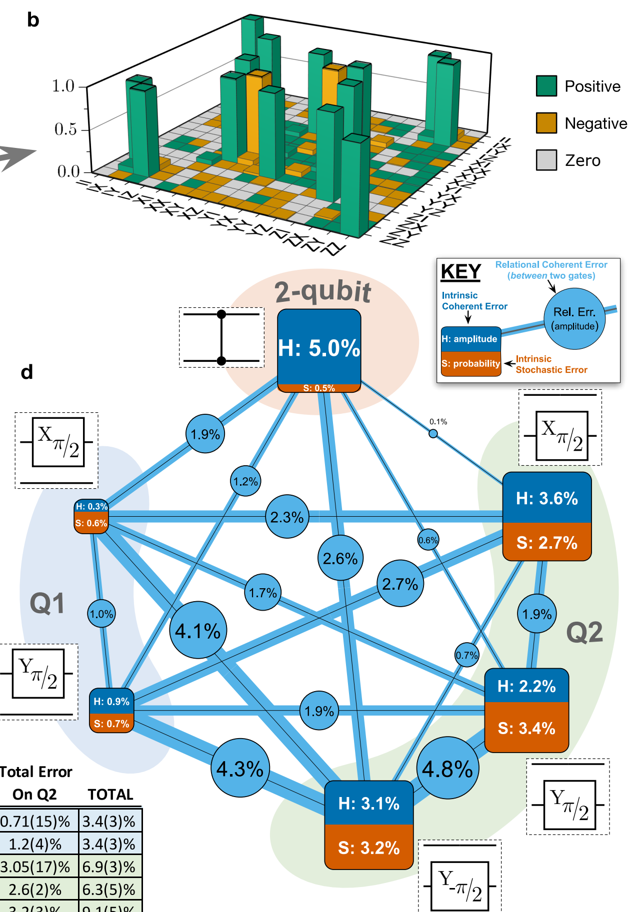
State	Fidelity	Concurrence
Φ^+	94.2(2.7)	0.91(5)
Φ^-	93.9(5.3)	0.89(10)
Ψ^+	94.1(2.3)	0.93(4)
Ψ^-	91.5(2.3)	0.87(4)

Fig. 2 | Tomography of nuclear Bell states. a, Each of the four Bell states has been generated using the same quantum circuit, only varying the initial spin state. **b-e,** Quantum state tomography results for (b) Φ^+; (c) Φ^-; (d) Ψ^+; (e) Ψ^- Bell state. No corrections have been applied to compensate readout errors. Hollow, black boxes indicate the outcome of an ideal measurement for each Bell state. **f,** Table of Bell state fidelities and concurrences. The error bars are estimated using Monte Carlo bootstrap re-sampling and represent 1σ confidence level.

erates a coherent ZZ rotation); Pauli-stochastic (S), also indexed by a single Pauli, cause probabilistic Pauli errors (e.g. S_{IX} causes probabilistic X errors on Q2); Pauli-correlation (C), and active (A), indexed by two Paulis, describe more exotic errors (see Methods) that were not detected in this experiment. We found that each gate's behavior could be described using just 13-14 elementary EGs: 3 local S errors and 3 local H errors acting on each of Q1 and Q2, and 1-2 entangling H errors (discussed in detail below). Extended Data Figure 8 shows those errors' rates, along with the process matrices and full EGs used to derive them. To get a higher-level picture of gate quality, we aggregate the rates of related errors (see Methods) to report total rates of stochastic and coherent errors on each qubit and on the entire 2-qubit system. We present two overall figures of merit in Figure 3a,c: generator infidelity and total error. Generator infidelity is closely related to entanglement infidelity, which accurately predicts average gate performance in realistic large-scale quantum processors and can be compared to fault-tolerance thresholds (see Methods and Supplementary Information S9). Total error is related to diamond norm (see Supplementary Information S9) and estimates worst-case gate performance in any circuit, including structured or periodic circuits. In Fig. 3c, we additionally report each gate's average gate fidelity on its target to ease comparison of these results with those from the literature.

The process matrices estimated by GST are not unique. An equivalent representation of the gate set can be constructed by a *gauge transformation* [34, 5] in which all process matrices are conjugated by some invertible matrix, $G_i \to MG_iM^{-1}$. Some gate errors, such as over/under-rotations or errors on idle spectator qubits, are nearly unaffected by choice of gauge; they are *intrinsic* to that gate. But other errors, such as a tilted rotation axis, can be shifted from one gate to another by changing gauge. These *relational* errors cannot be objectively associated with any particular gate. Recognizing this, we divide coherent errors into intrinsic and relational components (Fig. 3a,c). Intrinsic errors perturb a gate's eigenvalues, whereas relational errors perturb its eigenvectors. In Fig. 3a,c we follow standard GST practice by choosing a gauge that makes the gates as close to their targets as possible. This associates relational errors with individual gates, in a way that depends critically on the choice of gauge. But the magnitude of a given relational error between a set of gates is gauge-invariant, and Fig. 3d illustrates the total relational error between each pair of gates. In this work, we found evidence only for pairwise relational er-





c

Gate	On-target Average Gate Fidelity	Generator Infidelity On Q1	Generator Infidelity On Q2	Generator Infidelity TOTAL	Stochastic Error	*Intrinsic* Coherent Error	*Relational* Coherent Error	Total Error On Q1	Total Error On Q2	Total Error TOTAL
$X_{\frac{\pi}{2}} \otimes 1\!\!1$	**99.87(2)%**	**0.17(3)%**	0.47(5)%	0.68(6)%	0.60(6)%	0.34(14)%	2.8(3)%	2.0(3)%	0.71(15)%	3.4(3)%
$Y_{\frac{\pi}{2}} \otimes 1\!\!1$	**99.95(2)%**	**0.07(3)%**	0.61(6)%	0.75(6)%	0.67(6)%	0.9(3)%	2.6(3)%	1.0(2)%	1.2(4)%	3.4(3)%
$1\!\!1 \otimes X_{\frac{\pi}{2}}$	**99.48(3)%**	2.14(5)%	**0.69(4)%**	2.86(7)%	2.69(6)%	3.6(3)%	2.2(3)%	5.0(3)%	3.05(17)%	6.9(3)%
$1\!\!1 \otimes Y_{\frac{\pi}{2}}$	**99.41(5)%**	2.63(8)%	**0.79(6)%**	3.44(10)%	3.36(10)%	2.3(4)%	1.9(4)%	4.1(6)%	2.6(2)%	6.3(5)%
$1\!\!1 \otimes Y_{-\frac{\pi}{2}}$	**99.49(5)%**	2.60(13)%	**0.68(7)%**	3.54(16)%	3.18(15)%	3.1(4)%	5.1(5)%	4.2(7)%	3.2(3)%	9.1(5)%
CZ	**99.37(11)%**	0.11(10)%	0.66(11)%	**0.79(14)%**	0.54(13)%	4.9(3)%	0.7(3)%	3.2(4)%	4.1(3)%	5.5(4)%

Fig. 3 | Precise tomographic characterization of 1- and 2-qubit gate quality. Process matrices for all 6 gates (e.g., the CZ gate shown in **b**) were estimated using gate set tomography (GST) and represented as error generators with associated rates. **a**, Each gate's total error rate (columns) can be partitioned into coherent (blue) and stochastic (orange) components, then further into components acting on Q1 (left), Q2 (right), and on both at once (wide). Coherent errors are further partitioned into intrinsic (dark) and relational (light), which were assigned to specific gates by fixing a gauge. Each gate's generator infidelity (see Supplementary Information S9) is shown, on the whole 2-qubit system (hollow pins) and on its target qubit[s] only (black pins). The CZ gate's total infidelity is only 0.79(14)%. Single-qubit gates have on-target infidelities of 0.07(3)-0.79(6)%, but display significant crosstalk errors on the spectator qubit and unexpected entangling coherent (ZZ) errors. **c**, Error metrics for each gate are aggregated by type (stochastic/coherent) and support (Q1/Q2/total). In addition to generator infidelity, each gate's average gate fidelity on its target qubit[s] is shown, to facilitate comparison with literature. **d**, A gauge-invariant representation of relational errors between gates (e.g. misalignment of rotation axes) that were assigned to individual gates in **a**,**c** by fixing a gauge. Each gate is labeled with its intrinsic coherent (H) and stochastic (S) errors, while edges between two gates show the total amplitude of relational coherent error (misaligment) between them. Large gauge-invariant relational errors between single-qubit gates confirm that the entangling coherent errors observed in **a** are not an artifact of gauge-fixing.

rors, although more complex multi-gate relational errors are possible.

All 6 gates achieved on-target fidelities $> 99\%$, with infidelities as low as $0.07(3)\%$ on Q1 and $0.68(7)\%$ on Q2. However, we observed significant crosstalk on the spectator qubit during 1-qubit gates, resulting in full logic operations (1-qubit gate and spectator idle operation in parallel) with higher infidelities of $0.68(6)\%{-}3.5(2)\%$. Remarkably, the CZ gate's infidelity of $0.79(14)\%$ is almost on par with the single-qubit gates – a rare scenario in multi-qubit systems (Fig. 3a,c).

SPAM errors were estimated by GST as $1.05(4)\%$ on average, and as low as $0.25(3)\%$ for the $|\Uparrow\Uparrow\rangle$ state (Extended Data Figure 10). This is a unique feature of nuclear spin qubits, afforded by the quantum nondemolition nature of the measurement process [21] (Methods and Extended Data Fig. 5).

GST provided unambiguous evidence for a surprising relational error: weight-2 (entangling) H_{ZZ} and/or $H_{\mathbb{G}_i[ZZ]}$ coherent errors on each 1-qubit gate G_i, with amplitudes from $1.8 - 5.0\%$ (Extended Data Figure 8). These errors are consistent with an intermittent ZZ Hamiltonian during the gate pulses. After ruling out a wide range of possible error channels, we propose that the observed H_{ZZ} error arises from the spurious accumulation of geometric phase by the electron spin, caused by off-resonance leakage of microwave power near the ESR frequencies (Supplementary Information S9). This observation illustrates the diagnostic power of GST, which revealed an error channel we had not anticipated. It also shows GST's ability to unveil correlated and entangling errors, whose detection and prevention is of key importance for the realization of fault-tolerant quantum computers [35].

Three-qubit entanglement

The nuclear logic gates shown above would not scale beyond a single, highly localized cluster of donors. However, adding the hyperfine-coupled electron qubit yields a scalable heterogeneous architecture. Electron qubits decohere faster, but admit faster control. If high-fidelity entanglement between electron and nuclear qubits can be created, electron qubits can enable fast coherent communication between distant nuclei (via electron-electron entanglement, or physical shuttling) or serve as high-fidelity ancilla qubits for quantum error correction. To demonstrate this capability, we produce the maximally entangled three-qubit Greenberger-Horne-Zeilinger (GHZ) state $|\psi_{\mathrm{GHZ}}\rangle = (|\Uparrow\Uparrow\uparrow\rangle + |\Downarrow\Downarrow\downarrow\rangle)/\sqrt{2}$ using the pulse sequence shown in Fig. 4a. Starting from $|\Downarrow\Downarrow\downarrow\rangle$, an NMR $Y_{\pi/2}$ pulse at $\nu_{Q2|\downarrow}$ creates a coherent superposition state of nucleus 2, followed by a nuclear zCNOT gate (as in Fig. 2a) to produce a nuclear $|\Phi^+\rangle$ state, and an ESR X_π pulse at $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ to arrive at $|\psi_{\mathrm{GHZ}}\rangle$. Since the ESR frequency directly depends on the state of both nuclei, the latter pulse constitutes a natural 3-qubit Toffoli gate, making the creation of 3-qubit entanglement particularly simple, as in nitrogen-vacancy centres in diamond [36]. Executing Toffoli gates on electrons in quantum dots [37] requires more complex protocols, but can be simplified by a combination of exchange and microwave pulses [38].

Measuring the populations of the eight electron-nuclear states (Supplementary Information S7) after each step confirms the expected evolution from $|\Downarrow\Downarrow\downarrow\rangle$ to $|\psi_{\mathrm{GHZ}}\rangle$ (Fig. 4b). The evolution can be undone by applying the sequence in reverse, yielding a return probability to $|\Downarrow\Downarrow\downarrow\rangle$ of $89.6(9)\%$, including SPAM errors. As in the two-qubit case, measuring the populations is a useful sanity check but does not prove multipartite entanglement, which requires knowing the off-diagonal terms of the density matrix $\rho_{\mathrm{GHZ}} = |\psi_{\mathrm{GHZ}}\rangle\langle\psi_{\mathrm{GHZ}}|$.

Standard tomography methods require measuring the target state in different bases, obtained by rotating the qubits prior to measurement. However, the superposition of $|\Downarrow\Downarrow\downarrow\rangle$ and $|\Uparrow\Uparrow\uparrow\rangle$ dephases at a rate dominated by the electron dephasing time $T_{2\mathrm{e}}^* \approx 100$ μs (Extended Data Fig. 3), which is only marginally longer than the nuclear spin operation time $\approx 10{-}20$ μs. Therefore, the GHZ state will have significantly dephased by the time it is projected onto each measurement basis.

We circumvent this problem by adopting a tomography method that minimises the time spent in the GHZ state. An extension of a method first introduced for the measurement of electron-nuclear entanglement in spin ensembles [39], it is related to the parity scan commonly used in trapped ions [40] and superconducting circuits [41]. We repeat the reversal of the GHZ state (Fig. 4b) $N = 100$ times, each time introducing phase shifts $\theta_{1,2,3}$ to the rotation axes of the three reversal pulses, with $\theta_1 = 3\theta_2 = 9\theta_3 = 9N/125$. The return probability to $|\Downarrow\Downarrow\downarrow\rangle$ oscillates with N; the amplitude and phase of the oscillations yield the off-diagonal matrix element $\langle\Downarrow\Downarrow\downarrow|\,\rho_{\mathrm{GHZ}}\,|\Uparrow\Uparrow\uparrow\rangle = \rho_{18}$.

Since the ideal ρ_{GHZ} has nonzero elements only on its four corners, the populations ρ_{11}, ρ_{88} and the coherence ρ_{18} are sufficient to determine the GHZ state fidelity $\mathcal{F}_{\mathrm{GHZ}} = 92.5(1.0)\%$. Also here, SPAM errors remain included in total infidelity. By comparison, an 88% GHZ state fidelity has been reported in a triple quantum dot after removing SPAM errors, whereas the uncorrected fidelity is 45.8% [37]. This highlights the drastic effect of

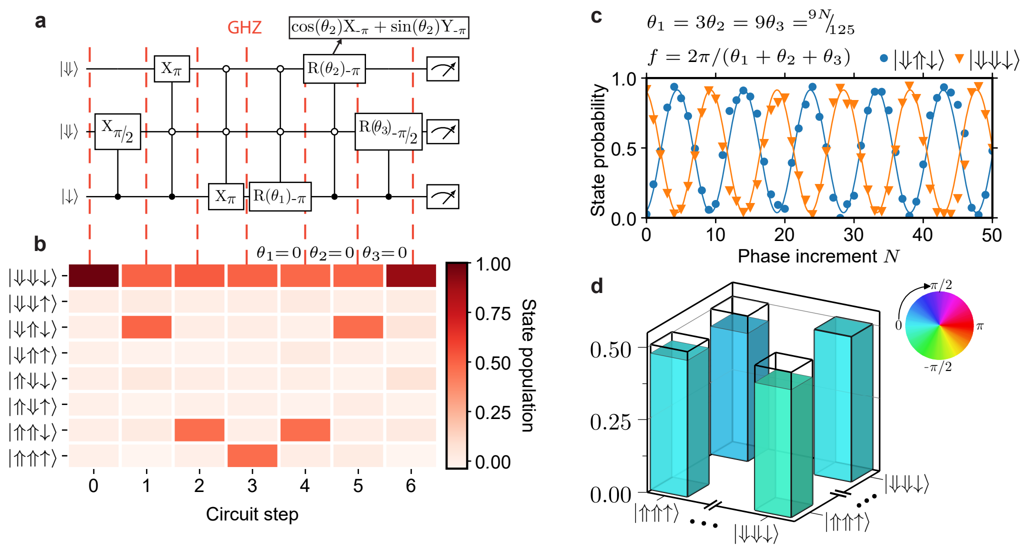


Fig. 4 | Creation and tomography of an electron-nuclear three-qubit GHZ state. a, Starting from $|{\Downarrow}{\Downarrow}{\downarrow}\rangle$, the first three gates generate an entangled three-qubit GHZ state. All eight state populations are read out (**b**) at each circuit step (red dashed lines), and estimated without correcting for SPAM errors (Supplementary Information S7). The final three gates $R(\theta_i)_\phi$ reverse the operations of the first three if the rotation angles are $\theta_1 = \theta_2 = \theta_3 = 0$, returning to the initial state in the absence of errors. The two gates that are conditional on Q_2 are composed of multiple pulses (Supplementary Information S6). **c,** The coherence between the GHZ components $|{\Downarrow}{\Downarrow}{\downarrow}\rangle$ and $|{\Uparrow}{\Uparrow}{\uparrow}\rangle$ is probed by incrementing the phases θ_i of the reversal pulses. This induces oscillations at frequency $f = 2\pi/(\theta_1+\theta_2+\theta_3)$ whose amplitude and phase correspond to the purity and phase relation between $|{\Downarrow}{\Downarrow}{\downarrow}\rangle$ and $|{\Uparrow}{\Uparrow}{\uparrow}\rangle$. **d,** Density-matrix extrema of the GHZ state. The state populations of the GHZ components $|{\Downarrow}{\Downarrow}{\downarrow}\rangle$ and $|{\Uparrow}{\Uparrow}{\uparrow}\rangle$ at circuit step 3 (**b**) provide the diagonal entries, while the oscillation amplitude and phase (**c**) provide the off-diagonal entries. From these values, the fidelity to the nearest GHZ state is estimated as 92.5(1.0)%, including SPAM.

SPAM of multi-qubit entanglement, and the robustness of our system against such errors. The different coherence and operation timescales for electron and nuclei need not be an obstacle for the use of such entangled states in scaled-up architectures, because all further entangling or shuttling operations between electrons will occur on $\simeq 1\ \mu$s time scales.

Outlook

The demonstration of 1-qubit, 2-qubit and SPAM errors at or below the 1% level highlight the potential of nuclear spins in silicon as a credible platform for fault-tolerant quantum computing. An often-quoted example, based on surface code quantum error correction, sets a fault-tolerance threshold of 0.56% for the entanglement infidelity of 1- and 2-qubit gates and the SPAM errors [6].

Several avenues are available to harness the high-fidelity operations demonstrated here. Replacing the ^{31}P donors with the higher-spin group-V analogues such as ^{123}Sb ($I = 7/2$) or ^{209}Bi ($I = 9/2$) would provide access to a much larger Hilbert space in which to encode quantum information. For example, a cluster of two ^{123}Sb donors contains the equivalent of six qubits in the nuclear spins, plus an electron qubit. An error-correcting code can be efficiently implemented in high-spin nuclei [42], where our method would provide a pathway for universal operations between the logical qubits encoded in each nucleus.

Moving to heavier group-V donors also allows the electrical control of the nuclear spins [43]. Combined with the electrical drive of the electron-nuclear 'flip-flop' transition [44], this implies the ability to control electron and nuclei by purely electrical means. In a two-donor system as shown here, the entangling CZ gate could similarly be obtained by an electrical 2π-pulse on a flip-flop transition.

The electron-nuclear entanglement we have demonstrated can be harnessed to scale up beyond a pair of nuclei coupled to the same electron. Neighbouring donor electrons can be entangled via exchange interaction by performing controlled-rotation resonant gates [9] or $\sqrt{\mathrm{SWAP}}$

gates [8]. Wider distances could be afforded by physically shuttling the electron across lithographic quantum dots [45, 46], while preserving the quantum information encoded in it [11]. Our methods would apply equally to isoelectronic nuclear spin centres like ^{73}Ge and ^{29}Si, where it has been shown that the nuclear qubit coherence is preserved while shuttling the electron across neighbouring dots [10]. Furthermore, electron spins can mediate the coherent interaction between nuclear spin qubits and microwave photons [47, 48]. Recent experiments on electron spin qubits in silicon report 1- and 2-qubit gate fidelities above 99% [49, 50]. Therefore, the fidelity of electron qubit operations will not constitute a bottleneck for the performance of electron-nuclear quantum processors. These examples illustrate the significance of universal high-fidelity two-qubit operations with nuclear spins in a platform like silicon, which can simultaneously host nuclear and electron spin qubits, lithographic quantum dots, and dense readout and control devices [19].

Methods

Device fabrication

The quantum processor is fabricated using methods compatible with standard silicon MOS processes. We start from a high quality silicon substrate (p-type ⟨100⟩; 10-20 Ωcm), on top of which a 900 nm thick epilayer of isotopically enriched ^{28}Si has been grown using low-pressure chemical vapour deposition (LPCVD). The residual ^{29}Si concentration is 730 ppm. Heavily-doped n$^+$ regions for Ohmic contacts and lightly-doped p regions for leakage prevention are defined by thermal diffusion of phosphorus and boron, respectively. A 200 nm thick SiO_2 field oxide is grown in a wet oxidation furnace. In the centre of the device, an opening of 20 μm × 40 μm is etched in the field oxide using HF acid. Immediately after, a 8 nm thick, high quality dry SiO_2 gate oxide is grown in this opening. In preparation for ion implantation, a 90 nm × 100 nm aperture is opened in a PMMA mask using electron-beam-lithography (EBL). The samples are implanted with P^+ ions at an acceleration voltage of 10 keV per ion. During implantation the samples were tilted by 8 degrees and the fluence was set at $1.4 \times 10^{12}/\mathrm{cm}^2$. Donor activation and implantation damage repair is achieved through the process of a rapid thermal annealing (5 seconds at 1000 °C). The gate layout is patterned around the implantation region in three EBL steps, each followed by aluminium thermal deposition (25 nm thickness for layer 1; 50 nm for layer 2; 100 nm for layer 3). Immediately after each metal deposition, the sample is exposed to a pure, low pressure (100 mTorr) oxygen atmosphere to form an Al_2O_3 layer, which electrically insulated the overlapping metal gates. At the last step, samples are annealed in a forming gas (400 °C, 15 min, 95% N_2 / 5% H_2) aimed at passivating the interface traps.

Experimental setup

The device was wire-bonded to a gold-plated printed circuit board and placed in a copper enclosure. The enclosure was placed in a permanent magnet array [51], producing a static magnetic field of 1.33 T at the device (see Extended Data Fig. 1 for field orientation). The board was mounted on a Bluefors BF-LD400 cryogen-free dilution refrigerator, reaching a base temperature of 14 mK, while the effective electron temperature was ≈ 150 mK.

DC bias voltages were applied to all gates using Stanford Research Systems (SRS) SIM928 voltage sources. A room-temperature resistive combiner was used for the fast donor gates (Extended Data Fig. 1) to add DC voltages to AC signals produced by the LeCroy Arbstudio 1104, which then passed through an 80 MHz low-pass filter; all other gates passed through a 20 Hz low-pass filter. All filtering takes place at the mixing chamber plate. The wiring includes graphite-coated flexible coaxial cables to reduce triboelectric noise [52].

Microwave pulses to induce ESR transitions were applied to an on-chip broadband antenna [53] using a Rohde & Schwarz SGS100A vector microwave source combined with an SGU100A upconverter. The microwave carrier frequency remained fixed at 37.1004125 GHz, while the output frequency was varied within a pulse sequence by mixing it with a radiofrequency (RF) signal using double-sideband modulation, i.e. by applying RF pulses to the in-phase port of the SGS100A IQ mixer (the quadrature port was terminated by a 50 Ω load). The carrier frequency was chosen such that whenever one sideband tone was resonant with an ESR pulse, the second sideband was off-resonant with all other ESR frequencies. To suppress microwave signals when not needed, 0 V was applied to the in-phase port of the IQ mixer. Under these circumstances, the carrier frequency is expected to be suppressed by 35 dB, according to the source data sheet. The RF pulses used for double-sideband modulation were generated by one of the two channels of the Agilent 81180A arbitrary waveform generator; the second channel delivered RF pulses to the microwave antenna to drive NMR transitions. The microwave signal for ESR and RF signal for NMR were combined in a Marki Microwave DPX-1721 diplexer.

The SET current passed through a Femto DLPCA-200 transimpedance amplifier (10^7 V/A gain, 50 kHz bandwidth), followed by an SRS SIM910 JFET post-amplifier (10^2 V/V gain), SRS SIM965 analog filter (50 kHz cutoff low-pass Bessel filter), and acquired via an AlazarTech ATS9440 PCI digitizer card. The instruments were triggered by a SpinCore PulseBlasterESR-PRO. The measurements instruments were controlled by Python code using the quantum measurement software packages QCoDeS and SilQ.

System Hamiltonian

The static Hamiltonian of our combined electron-nuclei system is

$$H_{\mathrm{s}} = -\gamma_{\mathrm{e}} B_0 \hat{S}_z - \gamma_{\mathrm{n}} B_0 (\hat{I}_{1,z} + \hat{I}_{2,z}) + A_1 \vec{S}\cdot\vec{I}_1 + A_2 \vec{S}\cdot\vec{I}_2, \quad (1)$$

where $\gamma_{\mathrm{e}} \approx -27.97$ GHz T^{-1} is the electron gyromagnetic ratio [54], $\gamma_{\mathrm{n}} \approx 17.23$ MHz T^{-1} is the nuclear gyromagnetic ratio [55], $\vec{S} = [\hat{S}_x, \hat{S}_y, \hat{S}_z]$ are the electron spin operators, and $\vec{I_i} = [\hat{I}_{i,x}, \hat{I}_{i,y}, \hat{I}_{i,z}]$ are the nuclear spin operators for nucleus $i \in 1, 2$. The static magnetic field $B_0 = 1.33$ T is aligned along $\hat{z}$, and $A_1 \approx 95$ MHz, ($A_2 \approx 9$ MHz) is the hyperfine interaction strength between the electron and nucleus 1 (2).

An AC drive applied to the microwave line is used to induce transitions between nuclear spin states and between electron spin states. The drive predominantly modulates the transverse magnetic field as

$$H_{\mathrm{rf}}(t) = -\gamma_{\mathrm{e}} \vec{B}_1 \cdot \vec{S} \sin\omega t - \gamma_{\mathrm{n}} \vec{B}_1 \cdot (\hat{I}_1 + \hat{I}_2) \sin\omega t, \quad (2)$$

where $\vec{B}_1$ is the oscillating magnetic field strength, primarily aligned along $\hat{y}$.

Electron spin readout

An electron spin readout is realized through the spin to charge conversion [56, 57]. This method utilizes a single electron transistor (SET) as both a charge sensor and an electron reservoir. The electron spin $|\downarrow\rangle$ and $|\uparrow\rangle$ states are separated by the Zeeman energy, which scales linearly with the external magnetic field. Thermal broadening of the SET at 100 mK is much smaller than the Zeeman splitting of two electron spin states. This means that, at the read position, the donor electron spin down state faces only occupied levels in the SET island (tunneling is prohibited) and the spin up state faces only unoccupied states and can freely tunnel out the SET island. This event will shift the energy ladder in the SET island, bringing it out of the Coulomb blockade, thus causing a burst in the current. This burst will last until $|\downarrow\rangle$ electron tunnels to the donor. If the electron has been projected to the $|\downarrow\rangle$ state then no change in the SET current will be recorded, as the electron cannot tunnel to the SET island. At the end of each read phase the electron spin is reinitialized in $|\downarrow\rangle$ for the next single shot cycle. The fidelity of single-shot electron readout and $|\downarrow\rangle$ initialisation by spin-dependent tunnelling is $\approx 80\%$ in this device. However, we further increase the initialisation fidelity by letting the electron thermalise to the lattice temperature for a time $\gg T_{1\mathrm{e}}$ (Fig. 3**b**) before triggering further operations.

Nuclear spin readout and initialisation

The readout of the two nuclear spin qubits is an extension of the well-known method developed for a single donor [21], based on the excitation of the electron bound to the nuclei, conditional on a particular nuclear state, followed by electron spin readout [20]. The same method is used to initialise the nuclei in a known state.

In the present system, consisting of an electron coupled to two ^{31}P donors with different hyperfine couplings $A_1 \gg A_2$, we find four well-separated electron spin resonance (ESR) frequencies (Fig. 1c), conditional on the $|\Downarrow\Downarrow\rangle, |\Downarrow\Uparrow\rangle, |\Uparrow\Downarrow\rangle, |\Uparrow\Uparrow\rangle$ nuclear states. An electron in the $|\downarrow\rangle$ state is initially drawn from a cold charge reservoir onto the donor cluster (independently of nuclear states). We then apply a microwave π-pulse at a particular ESR frequency, for instance $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ corresponding to the $|\Downarrow\Downarrow\rangle$ nuclear spin state, and then measure the electron spin. If it is found in the $|\uparrow\rangle$ state, then the nuclear spins are projected to the $|\Downarrow\Downarrow\rangle$ state. If the electron is $|\downarrow\rangle$ (i.e. the pulse at $\nu_{\mathrm{e}|\Downarrow\Downarrow}$ failed to flip it to $|\uparrow\rangle$), the nuclear spins are projected to the subspace orthogonal to the $|\Downarrow\Downarrow\rangle$ state. This constitutes a nuclear spins single-shot readout, with a fidelity given by the product of the electron single-shot readout fidelity (typically $\approx 80\%$) and the electron π-pulse fidelity ($\gg 99\%$).

This nuclear readout is a projective, approximately quantum non demolition (QND) process [21]. The ideal QND measurement relies on the observable I_z to commute with the Hamiltonian H_{int} describing an interaction between the observable and the measurement apparatus $[I_z, H_{\mathrm{int}}] = 0$ [58]. In our case the hyperfine terms $A_1 S_z I_{z1}$ and $A_2 S_z I_{z2}$ constitute H_{int}. The observation of nuclear spin quantum jumps originating from the electron measurement by spin-dependent tunnelling (ionization shock) hints at a deviation from QND nature of the readout process [21]. It implies the presence of terms of the form $A_{||}/2(S_+I_- + S_-I_+)$ in the hyperfine coupling, and possibly additional anisotropic terms, which do not commute with I_z. In our experiment, the deviation from the ideal QND measurement is extremely small, of order 10^{-6}, as shown in Extended Data Figure 5.

We exploit the near-perfect QND nature of the nuclear spin readout by repeating the cycle [load $|{\downarrow}\rangle$ – ESR π-pulse – electron readout] between 7 and 40 times, to substantially increase the nuclear single-shot readout fidelity. This is the fundamental reason why our average SPAM errors are $\approx 1\%$ (Extended Data Fig. 10), and we have thus reported Bell and GHZ state fidelities without removing SPAM errors from the estimate.

ESR and NMR calibration

Gate calibration

Both the 1-qubit NMR gates and the 2-qubit ESR gate were iteratively calibrated using a combination of GST and other tuning methods. Rabi flops were first used to obtain roughly calibrated 1-qubit NMR gates. Next, 1-qubit GST was repeatedly employed to identify and correct error contributions such as over-/under-rotations and detunings. Other routines such as the repeated application of gates were performed in between GST measurements to independently verify the improvements to 1-qubit gate fidelities of GST. The calibrated NMR $\pi/2$ pulse duration of Q1 (Q2) is 12.0 μs (25.3 μs). The discrepancy between the two durations is largely due to the hyperfine interaction enhancing the Rabi frequency of Q1 and reducing the Rabi frequency of Q2, combined with line reflections and filtering.

For the geometric 2-qubit gate based upon an electron 2π pulse, we found that a trivial calibration using Rabi flops already gave a near-optimal result. GST was then used for fine-tuning and for the detection of small error contributions such as a minor frequency shift. The calibrated ESR 2π pulse duration of the CZ gate is 1.89 μs at an output power of 20 dBm.

Periodic frequency recalibration

To keep the system tuned throughout the measurements, the NMR frequencies $\nu_{Q1|\downarrow}$ and $\nu_{Q2|\downarrow}$ and ESR frequency $\nu_{\mathrm{e}|\downarrow\downarrow}$ were calibrated every ten circuits. The ESR frequency was calibrated by measuring the ESR spectrum and selecting the frequency of the ESR peak. The NMR frequencies were measured by a variant of the Ramsey sequence, consisting of an $X_{\pi/2}$ and $Y_{\pi/2}$ separated by a wait time τ. An off-resonant RF pulse was applied during the wait time to mitigate any frequency shift caused by the absence of an RF drive. Since nuclear readout has a near-unity fidelity, this measurement should result in a nuclear flipping probability $P_{\text{flip}} = 0.5$ if the RF frequency f_{RF} matches the average NMR frequency f_{NMR} throughout the measurement. Therefore, any deviation of P_{flip} from 0.5 provides a direct estimate of the frequency mismatch $\Delta f = f_{\text{NMR}} - f_{\text{RF}} = \arcsin{(2P_{\text{flip}} - 1))}/(2\pi\tau)$, provided that $|\Delta f/\tau| < 0.25$. A higher τ more accurate estimates δf, while a lower τ results in the condition $|\Delta f/\tau| < 0.25$ being valid for a broader range of Δf. The NMR recalibration sequence iteratively increased the wait time $\tau = 40\ \mu\text{s} \rightarrow 100\ \mu\text{s} \rightarrow 160\ \mu\text{s}$ to ensure that the condition $|\Delta f/\tau| < 0.25$ remains satisfied while increasing the accuracy at which the NMR frequency is estimated. For each τ, the NMR frequency was estimated by repeating this sequence and updating the RF frequency until P_{flip} fell within the range [0.4, 0.6].

Measurement overhead

Instrument setups and calibration routines add a significant overhead to the GST measurements. An estimate of this overhead can be obtained by comparing the total measurement duration to the duration of a single pulse sequence. The 2Q GST measurement shown in Fig. 3 was acquired over 61 hours, during which 300-503 shots were acquired for each of the 1593 circuits. This results in an average duration of 340 ms per GST pulse sequence iteration. Compared to the average pulse sequence duration of around 121 ms, this corresponds to an overhead of 185%.

Effective mass theory simulations of the hyperfine interaction

To simulate the wave function of the third electron in the 2P system, the effective mass theory (EMT) model of the neutral 2P system in Ref. [59] is extended in a mean-field approach.

For short donor separations, the two inner electrons are tightly bound in a magnetically inactive singlet orbital. The third electron then only interacts with the inner ones to the extent that it experiences the Coulomb repulsion of their fixed charge distribution

$$V(\vec{r}) = \frac{e^2}{4\pi\epsilon_{\text{Si}}} \int \frac{\rho_{\text{S}}(\vec{r}')}{|\vec{r}' - \vec{r}|} d^3\vec{r}'. \tag{3}$$

Here, e is the electron charge, ϵ_{Si} the dielectric constant in silicon and $\rho_{\mathrm{S}}(\vec{r}')$ is the charge density of the tightly bound electrons found in Ref. [59]. The third electron is then effectively described by the sum of the 2P EMT Hamiltonian in an electric field [59] and the corresponding mean-field potential in Eq. (3).

Here, only 2P configurations along the [100] crystal axis with distances d≤7 nm and realistic fields E≤2 mV/nm are considered. In this regime the inter-donor exchange dominates the on-site exchange and the mean-field approach is justified.

The chosen basis is a combination of two STO-3G [59] orbitals, one variationally optimized at d=0.5 nm and the other at d=7 nm.

To compute the hyperfine interaction strength, the electron density at the nucleus is rescaled by a bunching factor of 440 [60]. The experimentally found hyperfine configuration is found for donors spaced 6.5 nm apart, and subjected to an electric field 2 mV/nm.

Gate set tomography experiments

We designed a customized GST experiment for a set of 6 logic gates: $X_{\pi/2}$ and $Y_{\pi/2}$ rotations on each qubit, an additional $Y_{-\pi/2}$ rotation on Q2, and the symmetric CZ gate between them. A basic 2-qubit GST experiment for this gate set comprises a list of quantum circuits defined by: (1) choosing a set of 75 short "germ" circuits that, when repeated, collectively amplify every error rate; (2) repeating each germ several times to times to form "germ power" circuits whose lengths are approximately $L = 1, 2, 4, \ldots L_{\max}$; and (3) prefacing and appending each germ power with each of 16 "preparation fiducial" circuits and each of 11 "measurement fiducial" circuits. We used $L_{\max} = 8$, yielding a set of 20606 circuits (this is not a simple multiplication because germ circuits with depth > 1 do not appear at shorter L). We eliminated 92% of these circuits using two techniques from [5]. First, we identified a subset of 18 germs that amplify any dominant errors in each gate (if $L_{\max}$ was very large, subdominant errors would get echoed away by dominant errors). This yielded a total of 50 germ powers. Second, for the $L > 1$ germ powers, we identified and eliminated pairs of fiducial circuits that provided redundant information. This trimmed the circuits per germ power from 176 to as few as 16, and the total number of circuits from 8800 to just 1592. Each of those circuits was repeated 300-500 times to gather statistics. We used maximum likelihood estimation (MLE) implemented in the `pyGSTi` software [61, 62] to estimate 16×16 2-qubit process matrices $\{G_i : \ i = 1 \ldots 6\}$ for all six operations.

Constructing and selecting reduced models

Process matrices are a comprehensive, but not especially transparent, representation of gate errors. So we used each gate's ideal target (unitary) operation $\mathbb{G}_i$ to construct an error generator [33] $\mathbb{L}_i = \log(G_i \mathbb{G}_i^{-1})$ that presents the same information more usefully. Representing noisy gates this way enables us to split each gate's total error into parts that act on Q1 only, Q2 only, or both qubits together – and then further into coherent and stochastic errors – to reveal those errors' sources and consequences. It also enables the construction of simple, efficient "reduced models" for gate errors, by identifying swaths of elementary error generators whose rates are indistinguishable from zero.

Pinning the coefficients of k elementary error generators to zero yields a reduced model with k fewer parameters, whose likelihood ($\mathcal{L}$) can be found by MLE. We evaluate the statistical significance of error rates that were pinned by seeing how much $\mathcal{L}$ declines. If a given error's true rate is zero, then pinning it to zero in the model reduces $2 \log \mathcal{L}$, on average, by 1 [63]. So when we pin k rates, we compute the "evidence ratio" $r = 2\Delta \log \mathcal{L}/k$, where $\Delta \log \mathcal{L}$ is the difference between the two models' likelihood [64]. If $r \leq 1$, the pinned rates are strictly negligible; if $r \leq 2$, then the smaller model is preferred by Akaike's information criterion (AIC) [65]; other criteria (e.g. the Bayesian BIC) impose higher thresholds. We used a slightly higher threshold and chose the smaller model whenever $r \leq 5$. Using this methodology, we constructed a model that describes the data well, in which just 83 (out of 1440) elementary errors' rates are significantly different from zero.

The rates of all the un-pinned elementary errors form a vector describing the noisy model. In general, un-physical gauge degrees of freedom [5] will give rise to a foliation of the model space into gauge manifolds on which the loglikelihood is constant. In our analysis, we work in the limit of small errors and gauge transformations where the space is approximately linear, and identify the subspace that is gauge invariant. We are able to construct a basis for the gauge-invariant subspace whose elements correspond to relational or intrinsic errors and have a definite type (H, S, or A), allowing us to decompose the model's total error as shown in Figure 3.

Extended Data Figure 8 presents each gate's 13-14 nonzero elementary error rates after projecting the error vector onto the gauge-invariant subspace (column 3), along with the process matrices (column 1) and error generators (column 2) from which they are derived. Here and elsewhere, error bars are 1σ confidence intervals computed using the Hessian of the loglikelihood function.

Aggregated error rates and metrics

Our GST analysis aims to identify specific gate errors and understand how these errors affect the overall performance of our system. It begins with the raw output of GST – rates of elementary errors on gates. We aggregate these error rates in different ways, yielding each gate's total error and infidelity, and partitioning those metrics into their components on Q1 or Q2 or both qubits together, in order to summarize different aspects of system performance. We additionally report average gate fidelities to facilitate comparison with the literature.

Gate errors by definition cause unintended changes in the state of the system. S error generators produce stochastic errors that transfer *probability* to erroneous states; H generators produce coherent errors that transfer *amplitude* to erroneous states. We can interpret the *rate* of an error generator, to first order, as the amount of erroneous probability (denoted ϵ for S generators) or amplitude (denoted θ for H generators) transferred by a single use of the gate when acting on one half of a maximally entangled state.

It is useful to group similar errors together and aggregate their rates. We classify and combine error generators according to:

- Their type (H or S),
- Their support (Q1, Q2, or joint),
- Whether they are intrinsic to a single gate, or relational between gates (H errors only; relational S errors were negligible).

The elementary error generators described in the main text have definite type and support. For example, the H_{XI} generator has type H and support on Q1. Any error generator on a given gate is intrinsic to that gate if it commutes with the gate, and relational otherwise. For example, if single-qubit $X_{\pi/2}$ and $Y_{\pi/2}$ gates produce rotations around axes that are separated by only 89° instead of 90°, then either gate can be considered perfect at the cost of assigning a 1° tilt error to the other gate. This error can be moved between the two gates by a gauge transformation M that rotates both gates by 1° around the Z-axis. This error is purely relational; it cannot be assigned definitively to one gate or the other, but can be unambiguously observed in circuits containing both gates.

To divide each gate's errors into intrinsic and relational components, we represent the gate's error generator as a vector in a space spanned by the H and S elementary error generators. Error generators that commute with the target gate form a subspace that is invariant under gauge transformations. The error generator's projection onto this space is its intrinsic component. Error generators in the complement of the intrinsic subspace are relational – they can be changed or eliminated by gauge transformations – and the projection of the gate's error generator onto this complement is its relational component.

To construct aggregated error metrics, we start by aggregating H and S rates separately. They add in different ways, because H error rates correspond to amplitudes while S error rates correspond to probabilities. Rates of S generators add directly ($\epsilon_{\text{agg}} = \sum_i \epsilon_i$), while rates of H generators add in quadrature ($\theta_{\text{agg}} = (\sum_i \theta_i^2)^{1/2}$). Combining H and S error rates into a single metric is trickier – there is no unique way to do so because the impact of coherent errors depends on how they interfere over the course of a circuit. We therefore consider two quantities: *total error* $\epsilon_{\text{tot}} = \epsilon_{\text{agg}} + \theta_{\text{agg}}$ and *generator infidelity* $\hat{\epsilon} = \epsilon_{\text{agg}} + \theta_{\text{agg}}^2$. Total error approximates the maximal rate at which gate errors could add up in any circuit, while infidelity quantifies the same errors' average impact in a random circuit.

Both of these metrics appear in Fig. 3, where in panels a, c, and d we report aggregated error rates that partition the overall error in various ways (see the discussion in S10 of the Supplement). We report a third metric, the *average gate fidelity* (AGF) on each gate's target qubit[s], in Fig. 3c and in the abstract to aid comparison with other published results. The on-target AGF provides an overall (and gauge-dependent) measure of the average performance of a gate when acting only on the target qubit(s). For a gate targeting Q1, it is defined as:

$$\bar{\epsilon}^{(Q1)} = 1 - \frac{1}{2}\int d\psi\ \langle\psi|\mathrm{tr}_{Q2}\left[e^{\mathbb{L}}\left(|\psi\rangle\langle\psi| \otimes \mathbb{I}\right)\right]|\psi\rangle \qquad (4)$$

For a two-qubit gate, the on-target AGF is simply the AGF of the two-qubit operation:

$$\bar{\epsilon} = 1 - \int d\psi\ \langle\psi|e^{\mathbb{L}}(|\psi\rangle\langle\psi|)|\psi\rangle, \qquad (5)$$

In both cases, $d\psi$ is the Haar measure (over 1-qubit states in Eq. 4 and over 2-qubit states in Eq. 5) and $\mathbb{L}$ is the error generator of the gate. Although AGF is provided for comparison to the literature, it is not a good predictor of performance in general circuits (see Supplemental

Information S9), and when we use the unqualified term "fidelity", it always denotes generator fidelity, $\hat{\epsilon}$. Section S9 of the Supplement includes an extensive discussion of overall gate error metrics and their relationships.

Data availability

The experimental data that support the findings of this study are available in Figshare with the identifier doi.org/10.6084/m9.figshare.c.5471706.

Code availability

Multivalley effective mass theory calculations, some of the results of which are illustrated in Fig. 1b, were performed using a fork of the code first developed in the production of Ref. [60] that was extended to include multielectron interactions as reported in Ref. [59]. Requests for a license for and copy of this code will be directed to points of contact at Sandia National Laboratories and the University of New South Wales, through the corresponding author.

References

[1] Kane, B. E. A silicon-based nuclear spin quantum computer. *Nature* **393**, 133 (1998).

[2] Vandersypen, L. M. & Chuang, I. L. NMR techniques for quantum control and computation. *Reviews of Modern Physics* **76**, 1037 (2005).

[3] Saeedi, K. *et al.* Room-temperature quantum bit storage exceeding 39 minutes using ionized donors in silicon-28. *Science* **342**, 830 (2013).

[4] Filidou, V. *et al.* Ultrafast entangling gates between nuclear spins using photoexcited triplet states. *Nature Physics* **8**, 596–600 (2012).

[5] Nielsen, E. *et al.* Gate set tomography. *Quantum* **5**, 557 (2021).

[6] Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: Towards practical large-scale quantum computation. *Physical Review A* **86**, 032324 (2012).

[7] Harvey-Collard, P. *et al.* Coherent coupling between a quantum dot and a donor in silicon. *Nature Communications* **8**, 1–6 (2017).

[8] He, Y. *et al.* A two-qubit gate between phosphorus donor electrons in silicon. *Nature* **571**, 371–375 (2019).

[9] Mądzik, M. T. *et al.* Conditional quantum operation of two exchange-coupled single-donor spin qubits in a MOS-compatible silicon device. *Nature Communications* **12**, 181 (2021).

[10] Hensen, B. *et al.* A silicon quantum-dot-coupled nuclear spin qubit. *Nature Nanotechnology* **15**, 13–17 (2020).

[11] Yoneda, J. *et al.* Coherent spin qubit transport in silicon. *Nature Communications* **12**, 4114 (2021).

[12] Zhong, M. *et al.* Optically addressable nuclear spins in a solid with a six-hour coherence time. *Nature* **517**, 177–180 (2015).

[13] Muhonen, J. T. *et al.* Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking. *Journal of Physics: Condensed Matter* **27**, 154205 (2015).

[14] Bradley, C. *et al.* A ten-qubit solid-state spin register with quantum memory up to one minute. *Physical Review X* **9**, 031045 (2019).

[15] Bourassa, A. *et al.* Entanglement and control of single nuclear spins in isotopically engineered silicon carbide. *Nature Materials* **19**, 1319–1325 (2020).

[16] Waldherr, G. *et al.* Quantum error correction in a solid-state hybrid spin register. *Nature* **506**, 204 (2014).

[17] Bhaskar, M. K. *et al.* Experimental demonstration of memory-enhanced quantum communication. *Nature* **580**, 60–64 (2020).

[18] Pompili, M. *et al.* Realization of a multinode quantum network of remote solid-state qubits. *Science* **372**, 259–264 (2021).

[19] Vandersypen, L. *et al.* Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent. *npj Quantum Information* **3**, 1–10 (2017).

[20] Morello, A. *et al.* Single-shot readout of an electron spin in silicon. *Nature* **467**, 687–691 (2010).

[21] Pla, J. J. *et al.* High-fidelity readout and control of a nuclear spin qubit in silicon. *Nature* **496**, 334–338 (2013).

[22] Pla, J. J. *et al.* A single-atom electron spin qubit in silicon. *Nature* **489**, 541–545 (2012).

[23] Ivie, J. A. *et al.* The impact of stochastic incorporation on atomic-precision Si:P arrays. *arXiv preprint arXiv:2105.12074* (2021).

[24] Hile, S. J. *et al.* Addressable electron spin resonance using donors and donor molecules in silicon. *Science Advances* **4**, eaaq1459 (2018).

[25] Anandan, J. The geometric phase. *Nature* **360**, 307–313 (1992).

[26] James, D. F. V., Kwiat, P. G., Munro, W. J. & White, A. G. Measurement of qubits. *Physical Review A* **64**, 052312 (2001).

[27] Dehollain, J. P. *et al.* Optimization of a solid-state electron spin qubit using gate set tomography. *New Journal of Physics* **18**, 103018 (2016).

[28] Blume-Kohout, R. *et al.* Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature Communications* **8**, 14485 (2017).

[29] Huang, W. *et al.* Fidelity benchmarks for two-qubit gates in silicon. *Nature* **569**, 532–536 (2019).

[30] Xue, X. *et al.* Benchmarking gate fidelities in a Si/SiGe two-qubit device. *Physical Review X* **9**, 021011 (2019).

[31] Kimmel, S., da Silva, M. P., Ryan, C. A., Johnson, B. R. & Ohki, T. Robust extraction of tomographic information via randomized benchmarking. *Physical Review X* **4**, 011050 (2014).

[32] Carignan-Dugas, A., Wallman, J. J. & Emerson, J. Bounding the average gate fidelity of composite channels using the unitarity. *New Journal of Physics* **21**, 053016 (2019).

[33] Blume-Kohout, R. *et al.* A taxonomy of small markovian errors. *arXiv preprint arXiv:2103.01928* (2021).

[34] Proctor, T., Rudinger, K., Young, K., Sarovar, M. & Blume-Kohout, R. What randomized benchmarking actually measures. *Physical Review Letters* **119**, 130502 (2017).

[35] Novais, E. & Mucciolo, E. R. Surface code threshold in the presence of correlated errors. *Physical Review Letters* **110**, 010502 (2013).

[36] Neumann, P. *et al.* Multipartite entanglement among single spins in diamond. *Science* **320**, 1326–1329 (2008).

[37] Takeda, K. *et al.* Quantum tomography of an entangled three-qubit state in silicon. *Nature Nanotechnology* **16**, 965–969 (2021).

[38] Gullans, M. & Petta, J. Protocol for a resonantly driven three-qubit toffoli gate with silicon spin qubits. *Physical Review B* **100**, 085419 (2019).

[39] Mehring, M., Mende, J. & Scherer, W. Entanglement between an electron and a nuclear spin 1/2. *Physical Review Letters* **90**, 153001 (2003).

[40] Sackett, C. A. *et al.* Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000).

[41] Wei, K. X. *et al.* Verifying multipartite entangled greenberger-horne-zeilinger states via multiple quantum coherences. *Physical Review A* **101**, 032343 (2020).

[42] Gross, J. A., Godfrin, C., Blais, A. & Dupont-Ferrier, E. Hardware-efficient error-correcting codes for large nuclear spins. *arXiv preprint arXiv:2103.08548* (2021).

[43] Asaad, S. *et al.* Coherent electrical control of a single high-spin nucleus in silicon. *Nature* **579**, 205–209 (2020).

[44] Tosi, G. *et al.* Silicon quantum processor with robust long-distance qubit couplings. *Nature Communications* **8**, 450 (2017).

[45] Pica, G., Lovett, B. W., Bhatt, R. N., Schenkel, T. & Lyon, S. A. Surface code architecture for donors and dots in silicon with imprecise and nonuniform qubit couplings. *Physical Review B* **93**, 035306 (2016).

[46] Buonacorsi, B. *et al.* Network architecture for a topological quantum computer in silicon. *Quantum Science and Technology* **4**, 025003 (2019).

[47] Tosi, G., Mohiyaddin, F. A., Tenberg, S., Laucht, A. & Morello, A. Robust electric dipole transition at microwave frequencies for nuclear spin qubits in silicon. *Physical Review B* **98**, 075313 (2018).

[48] Mielke, J., Petta, J. R. & Burkard, G. Nuclear spin readout in a cavity-coupled hybrid quantum dot-donor system. *PRX Quantum* **2**, 020347 (2021).

[49] Xue, X. *et al.* Computing with spin qubits at the surface code error threshold. *arXiv preprint arXiv:2107.00628* (2021).

[50] Noiri, A. *et al.* Fast universal quantum control above the fault-tolerance threshold in silicon. *arXiv preprint arXiv:2108.02626* (2021).

[51] Adambukulam, C. *et al.* An ultra-stable 1.5 tesla permanent magnet assembly for qubit experiments at cryogenic temperatures. *arXiv preprint arXiv:2010.02455* (2020).

[52] Kalra, R. *et al.* Vibration-induced electrical noise in a cryogen-free dilution refrigerator: Characterization, mitigation, and impact on qubit coherence. *Review of Scientific Instruments* **87**, 073905 (2016).

[53] Dehollain, J. *et al.* Nanoscale broadband transmission lines for spin qubit control. *Nanotechnology* **24**, 015202 (2012).

[54] Feher, G. Electron spin resonance experiments on donors in silicon. i. electronic structure of donors by the electron nuclear double resonance technique. *Physical Review* **114**, 1219 (1959).

[55] Steger, M. *et al.* Optically-detected NMR of optically-hyperpolarized ^{31}P neutral donors in ^{28}Si. *Journal of Applied Physics* **109**, 102411 (2011).

[56] Elzerman, J. M. *et al.* Single-shot read-out of an individual electron spin in a quantum dot. *Nature* **430**, 431 (2004).

[57] Morello, A. *et al.* Architecture for high-sensitivity single-shot readout and control of the electron spin of individual donors in silicon. *Physical Review B* **80**, 081307 (2009).

[58] Braginsky, V. B. & Khalili, F. Y. Quantum nondemolition measurements: the route from toys to tools. *Reviews of Modern Physics* **68**, 1 (1996).

[59] Joecker, B. *et al.* Full configuration interaction simulations of exchange-coupled donors in silicon using multi-valley effective mass theory. *New Journal of Physics* (2021).

[60] Gamble, J. K. *et al.* Multivalley effective mass theory simulation of donors in silicon. *Physical Review B* **91**, 235318 (2015).

[61] Nielsen, E. *et al.* Python GST implementation (PyGSTi) v. 0.9. Tech. Rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States) (2019).

[62] Nielsen, E. *et al.* Probing quantum processor performance with pyGSTi. *Quantum Science and Technology* **5**, 044002 (2020).

[63] Wilks, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics* **9**, 60 – 62 (1938).

[64] Nielsen, E., Rudinger, K., Proctor, T., Young, K. & Blume-Kohout, R. Efficient flexible characterization of quantum processors with nested error models. *arXiv preprint arXiv:2103.02188* (2021).

[65] Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike*, 199–213 (Springer, 1998).

[66] Tenberg, S. B. *et al.* Electron spin relaxation of single phosphorus donors in metal-oxide-semiconductor nanoscale devices. *Physical Review B* **99**, 205306 (2019).

[67] Hsueh, Y.-L. *et al.* Spin-lattice relaxation times of single donors and donor clusters in silicon. *Physical review letters* **113**, 246406 (2014).

Acknowledgements

We acknowledge helpful conversations with W. Huang, R. Rahman, S. Seritan, and C. H. Yang and technical support from T. Botzem. The research was supported by the Australian Research Council (Grant no. CE170100012), the US Army Research Office (Contract no. W911NF-17-1-0200), and the Australian Department of Industry, Innovation and Science (Grant No. AUSMURI000002). We acknowledge support from the Australian National Fabrication Facility (ANFF). This material is based upon work supported in part by the iHPC facility at UTS, by the by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research's Quantum Testbed Pathfinder and Early Career Research Programs, and by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers (Quantum Systems Accelerator). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI, the U.S. Department of Energy, or the U.S. Government.

Author contributions

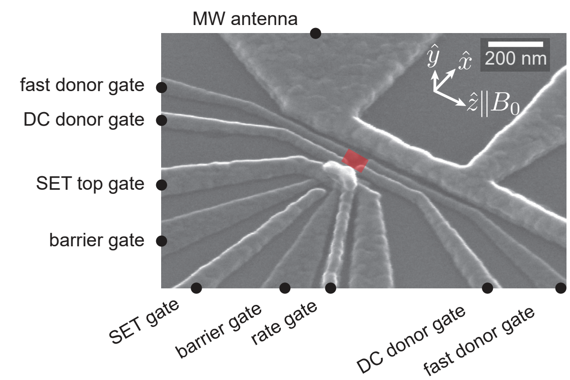
M.T.M., V.S. and F.E.H. fabricated the device, with A.M.'s and A.S.D.'s supervision, on an isotopically-enriched ^{28}Si wafer supplied by K.M.I.. A.M.J., B.C.J. and D.N.J. designed and performed the ion implantation. M.T.M. and S.A. performed the experiments and analysed the data, with A.L. and A.M.'s supervision. B.J. and A.D.B. developed and applied computational tools to calculate the electron wavefunction and the Hamiltonian evolution. A.Y. designed the initial GST sequences, with C.F.'s supervision. K.M.R., E.N., K.C.Y., T.J.P. and R.B.-K. developed the and applied the GST method. A.M., R.B-K., M.T.M. and S.A. wrote the manuscript, with input from all coauthors.

These authors contributed equally: Mateusz T. Mądzik, Serwan Asaad.

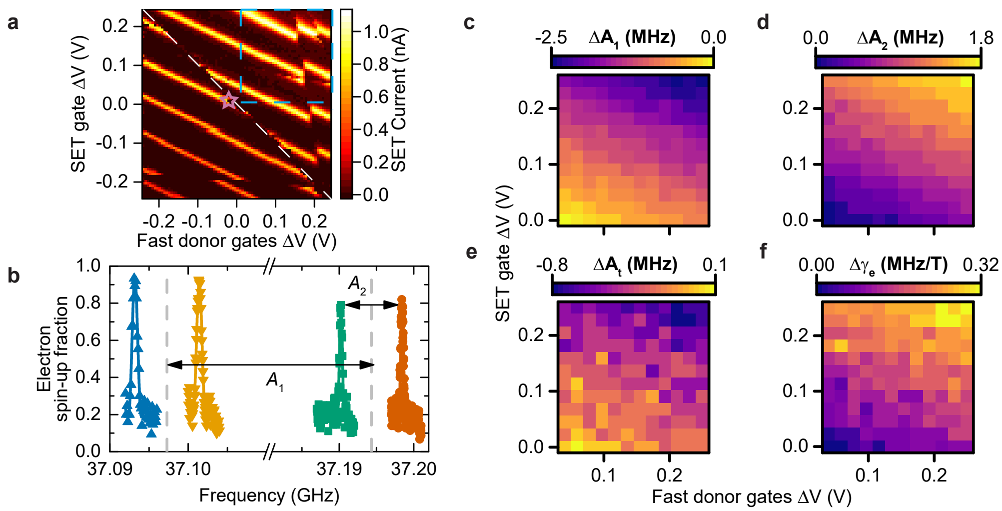
Correspondence and requests for materials should be addressed to Andrea Morello, a.morello@unsw.edu.au.

Competing interest The authors declare no competing interests.

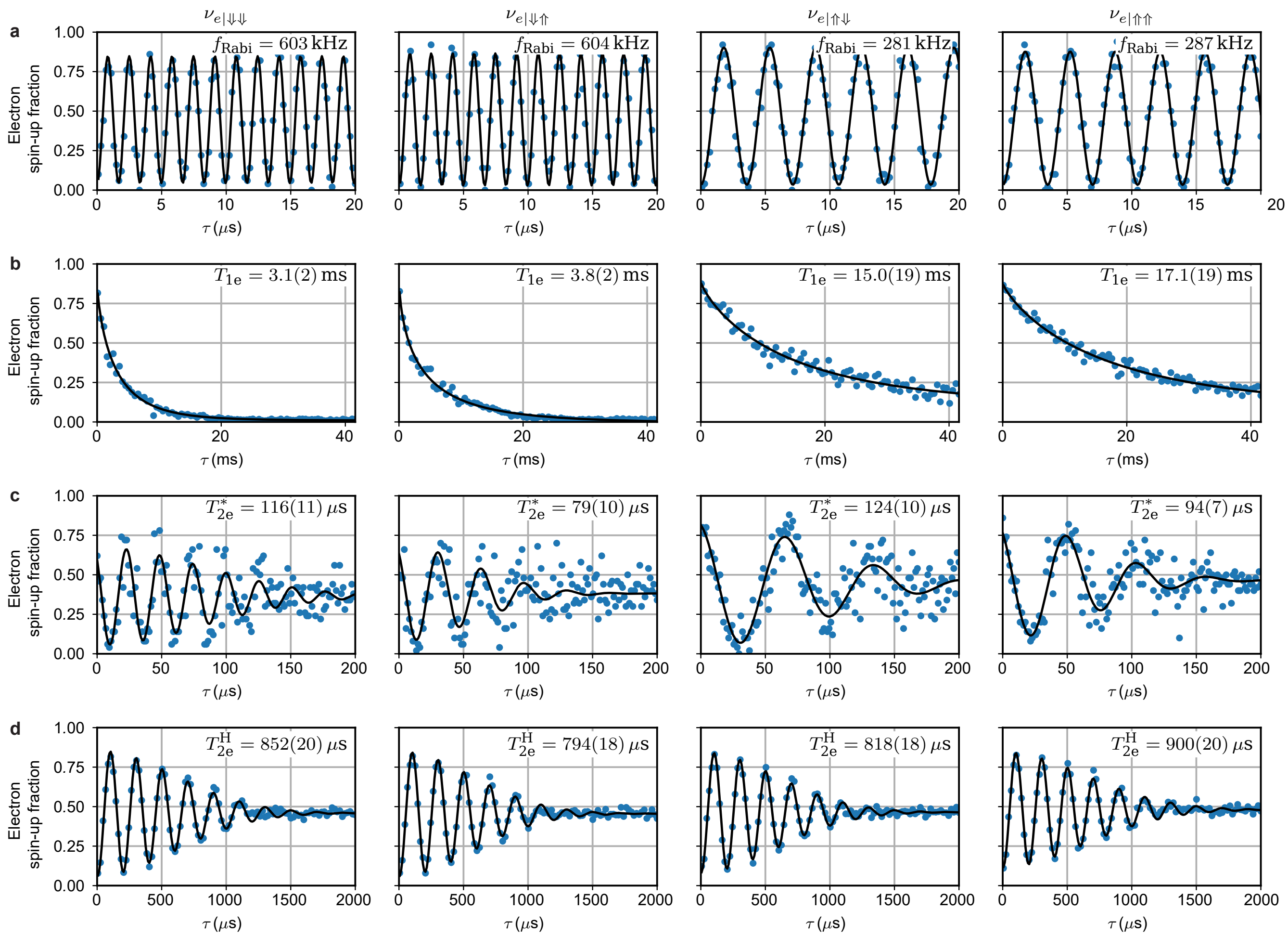
Extended data figures and tables



Extended Data Fig. 1 | Device layout. Scanning electron micrograph of a device identical to the one used in this experiment. ^{31}P donor atoms are implanted in the region marked by the orange rectangle, using a fluence of $1.4 \times 10^{12}/\mathrm{cm}^2$ which results in a most probably inter-donor spacing of approximately 8 nm. Four metallic gates are fabricated around the implantation region, and used to modify the electrochemical potential of the donors. A nearby SET, formed using the SET top gate and barrier gates, enables charge sensing of a single donor atom, as well as its electron spin through spin-to-charge conversion (Methods). The tunnel coupling between the donors and SET is tuned by the rate gate situated between the SET and donor implant region. A nearby microwave (MW) antenna is used for ESR and NMR of the donor electron and nuclear spins, respectively.



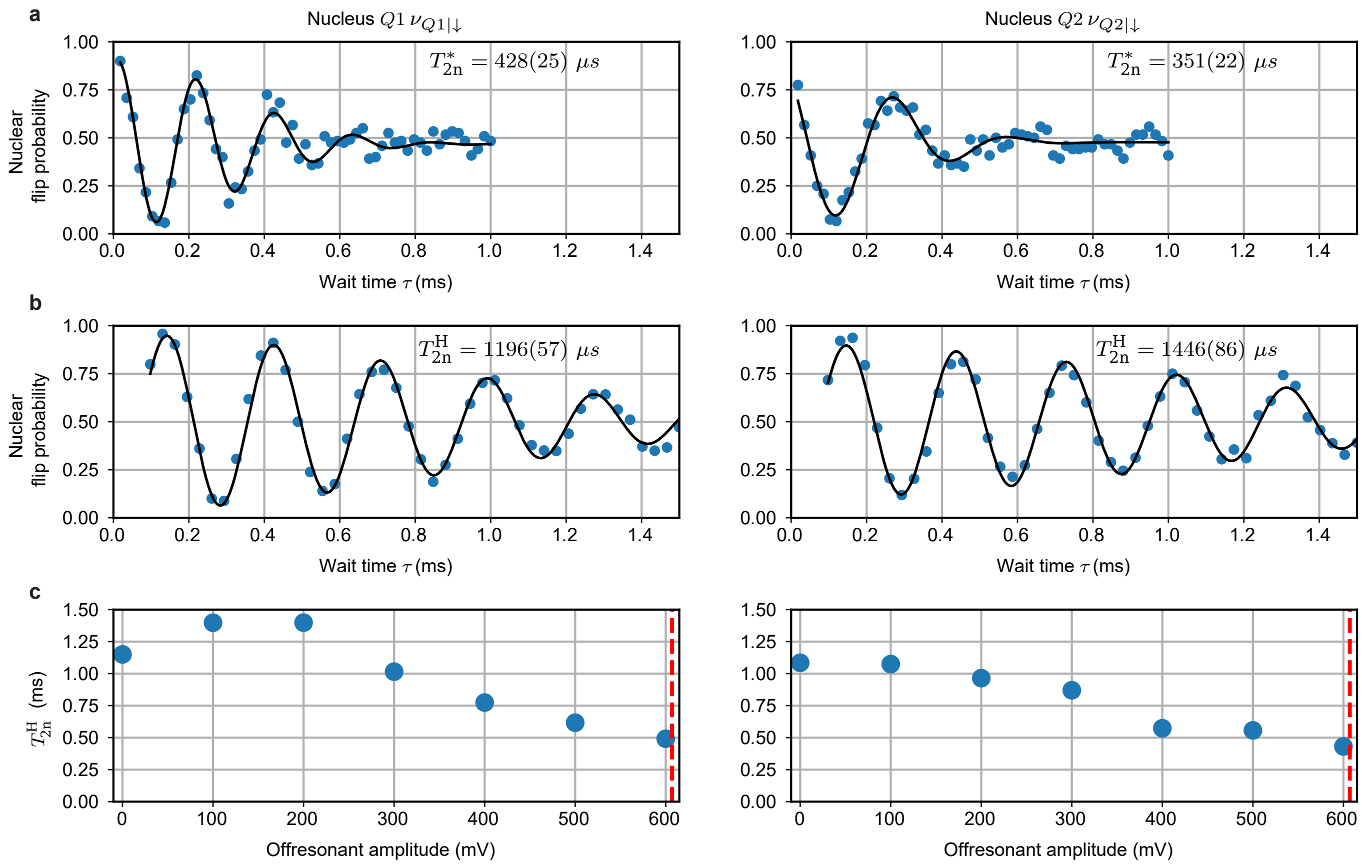
Extended Data Fig. 2 | Electrical tunability of the hyperfine interaction and the electron gyromagnetic ratio. a, Map of the SET current as a function of SET gate and fast donor gates (pulsed jointly). The white dashed line indicates the location in gate space where the 2P donor cluster changes its charge state. The third, hyperfine-coupled electron is present on the cluster in the region to the right of the line. Electron spin readout is performed at the location indicated by the pink star. **b,** ESR spectrum of the electron bound to the 2P cluster, acquired while the system was tuned within the blue dashed rectangle in panel **a**. The hyperfine couplings A_1, A_2 are extracted from ESR frequencies as shown, namely $A_1 = (\nu_{\mathrm{e}|\Uparrow\Downarrow} + \nu_{\mathrm{e}|\Uparrow\Uparrow})/2 - (\nu_{\mathrm{e}|\Downarrow\Downarrow} + \nu_{\mathrm{e}|\Downarrow\Uparrow})/2$; $A_2 = \nu_{\mathrm{e}|\Uparrow\Uparrow} - \nu_{\mathrm{e}|\Uparrow\Downarrow}$. **c-d,** Extracted hyperfine couplings within the marked area. The data shows that A_1 decreases and A_2 increases upon moving the operation point towards higher gate voltages and away from the donor readout position. **e,** A small change is also observed in the sum of the two hyperfine interactions $A_{\mathrm{t}} = A_1 + A_2$. **f,** Electrical modulation (Stark shift) of the electron gyromagnetic ratio γ_e, extracted from the shift of the average of the hyperfine-split electron resonances. The ESR frequencies can be tuned with fast donor gates at the rate of $\Delta\nu_{\mathrm{e}|\Uparrow\Uparrow} = 0.3$ MHz/V; $\Delta\nu_{\mathrm{e}|\Uparrow\Downarrow} = 5.2$ MHz/V; $\Delta\nu_{\mathrm{e}|\Downarrow\Uparrow} = 7.6$ MHz/V; $\Delta\nu_{\mathrm{e}|\Downarrow\Downarrow} = 2.4$ MHz/V.



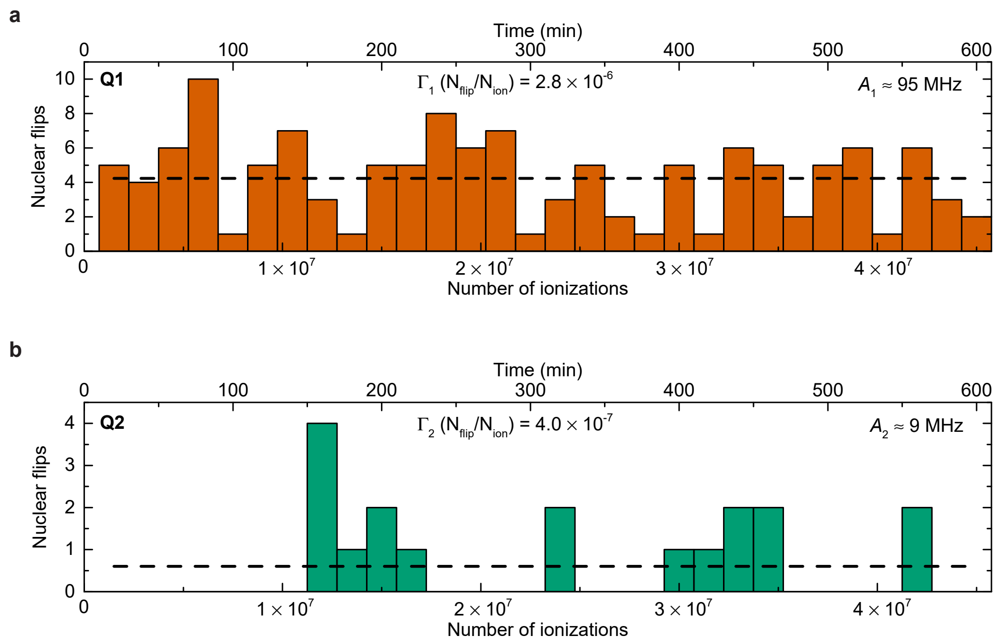
$\nu_{e|\Downarrow\Downarrow}$
$\nu_{e|\Downarrow\Uparrow}$
$\nu_{e|\Uparrow\Downarrow}$
$\nu_{e|\Uparrow\Uparrow}$
a
b
c
d
Electron spin-up fraction
$f_{\mathrm{Rabi}} = 603\,\mathrm{kHz}$
$f_{\mathrm{Rabi}} = 604\,\mathrm{kHz}$
$f_{\mathrm{Rabi}} = 281\,\mathrm{kHz}$
$f_{\mathrm{Rabi}} = 287\,\mathrm{kHz}$
$T_{1\mathrm{e}} = 3.1(2)\,\mathrm{ms}$
$T_{1\mathrm{e}} = 3.8(2)\,\mathrm{ms}$
$T_{1\mathrm{e}} = 15.0(19)\,\mathrm{ms}$
$T_{1\mathrm{e}} = 17.1(19)\,\mathrm{ms}$
$T_{2\mathrm{e}}^{*} = 116(11)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{*} = 79(10)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{*} = 124(10)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{*} = 94(7)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{\mathrm{H}} = 852(20)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{\mathrm{H}} = 794(18)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{\mathrm{H}} = 818(18)\,\mu\mathrm{s}$
$T_{2\mathrm{e}}^{\mathrm{H}} = 900(20)\,\mu\mathrm{s}$
$\tau\,(\mu\mathrm{s})$
$\tau\,(\mathrm{ms})$

1091

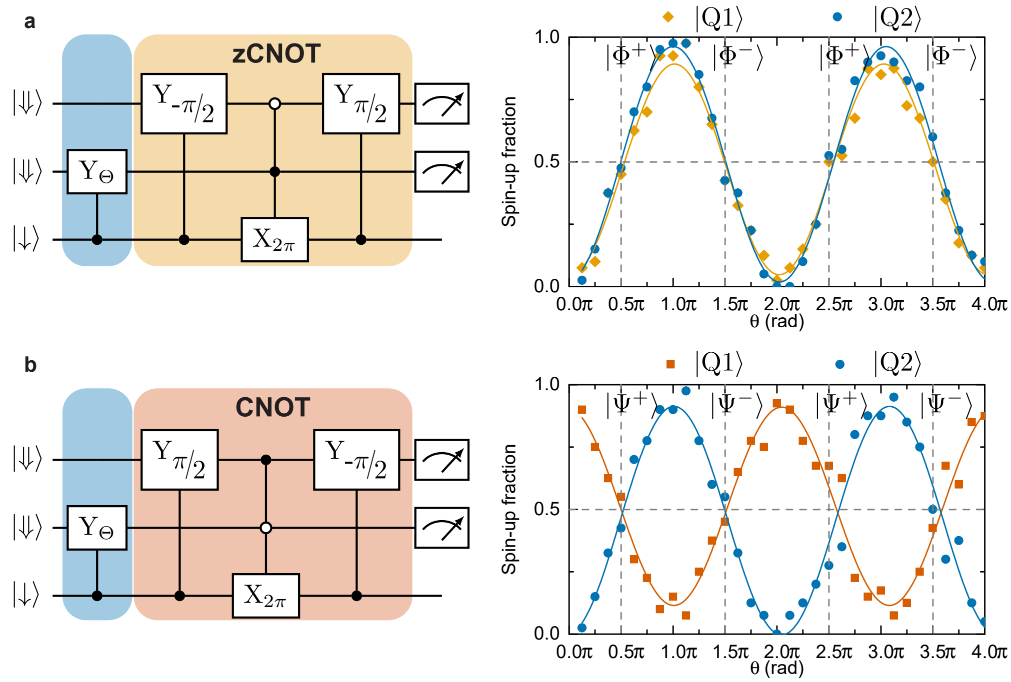
Extended Data Fig. 3 | Coherence metrics of the electron spin qubit. The columns correspond to the nuclear configurations $|\Downarrow\Downarrow\rangle$, $|\Downarrow\Uparrow\rangle$, $|\Uparrow\Downarrow\rangle$, $|\Uparrow\Uparrow\rangle$, respectively. All measurements start with the electron spin initialized in the $|\downarrow\rangle$ state. Error bars are 1σ confidence intervals. **a,** Electron Rabi oscillations. The measurements were performed by applying a resonant ESR pulse of increasing duration. The different Rabi frequencies f_{Rabi} on each resonance are likely due to a frequency-dependent response of the on-chip antenna and the cable connected to it. **b,** Electron spin-lattice relaxation times $T_{1\mathrm{e}}$. Measurements were obtained by first adiabatically inverting the electron spin to $|\uparrow\rangle$, followed by a varying wait time τ before electron readout. The observed relaxation times are nearly three orders of magnitude shorter than typically observed in single-electron, single-donor devices [66], and even shorter compared to 1e-2P clusters. This strongly suggests that the measured electron is the third one, on top of two more tightly-bound electrons which form a singlet spin state [67]. We also observe a strong dependence of $T_{1\mathrm{e}}$ on nuclear spin configuration. **c,** Electron dephasing times $T_{2\mathrm{e}}^{*}$. The measurements were conducted by performing a Ramsey experiment, i.e. by applying two $\pi/2$ pulses separated by a varying wait time τ, followed by electron readout. The Ramsey fringes are fitted to a function of the form $P_{\uparrow}(\tau) = C_0 + C_1 \cos(\Delta\omega \cdot \tau + \Delta\phi)\exp[-(\tau/T_{2\mathrm{e}}^{*})^2]$, where $\Delta\omega$ is the frequency detuning and $\Delta\phi$ is a phase offset. The observed $T_{2\mathrm{e}}^{*}$ times are comparable to previous values for electrons coupled to a single ^{31}P nucleus. **d,** Electron Hahn-echo coherence times $T_{2\mathrm{e}}^{\mathrm{H}}$, obtained by adding a π refocusing pulse to the Ramsey sequence. We also varied the phase of the final $\pi/2$ pulse at a rate of one period per $\tau = (5\ \mathrm{kHz})^{-1}$., to introduce oscillations in the spin-up fraction which help improve the fitting. The curves are fitted to the same function used to fit the Ramsey fringes, with fixed $\Delta\omega = 5$ kHz. The measured $T_{2\mathrm{e}}^{\mathrm{H}}$ times are similar to previous observations for electrons coupled to a single ^{31}P nucleus.



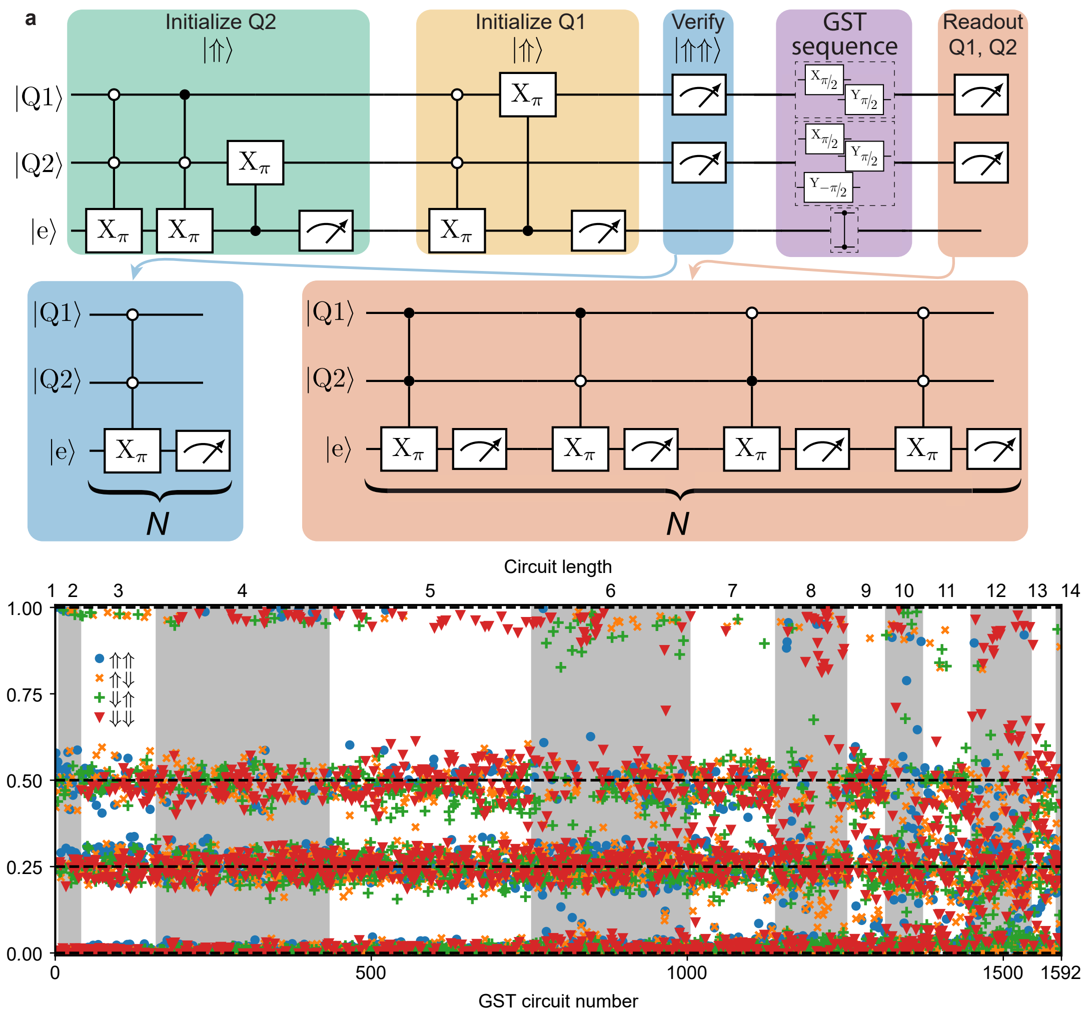
Extended Data Fig. 4 | Nuclear spin coherence times. Panels in column 1 (2) correspond to nucleus $Q1$ ($Q2$). Error bars are 1σ confidence intervals. **a,** Nuclear dephasing times T_{2n}^{*}, obtained from a Ramsey experiment. Results are fitted with a decaying sinusoid with fixed exponent factor 2 (see Extended Data Fig. 3). **b,** Nuclear Hahn-echo coherence times T_{2n}^{H}. To improve fitting, oscillations are induced by incrementing the phase of the final $\pi/2$ pulse with τ at a rate of one period per $(3.5\text{ kHz})^{-1}$. Results are fitted with a decaying sinusoid with fixed exponent factor 2 (see Extended Data Fig. 3). **c,** Dependence of T_{2n}^{H} on the amplitude of an off-resonance pulse. We perform this experiment to study whether a qubit, nominally left idle (or, in quantum information terms, subjected to an identity gate) is affected by the application of an RF pulse to the other qubit, at a vastly different frequency. Here, during the idle times between NMR pulses, an RF pulse is applied at a fixed frequency 20 MHz – far off-resonance from both qubits' transitions – with varying amplitude V_{RF}. The red dashed line indicates the applied RF amplitude for NMR pulses throughout the experiment. We observe a slow decrease of T_{2n}^{H} with increasing V_{RF}. This is qualitatively consistent with the observation of large stochastic errors on the idle qubit, as extracted by the GST analysis in Fig. 3.



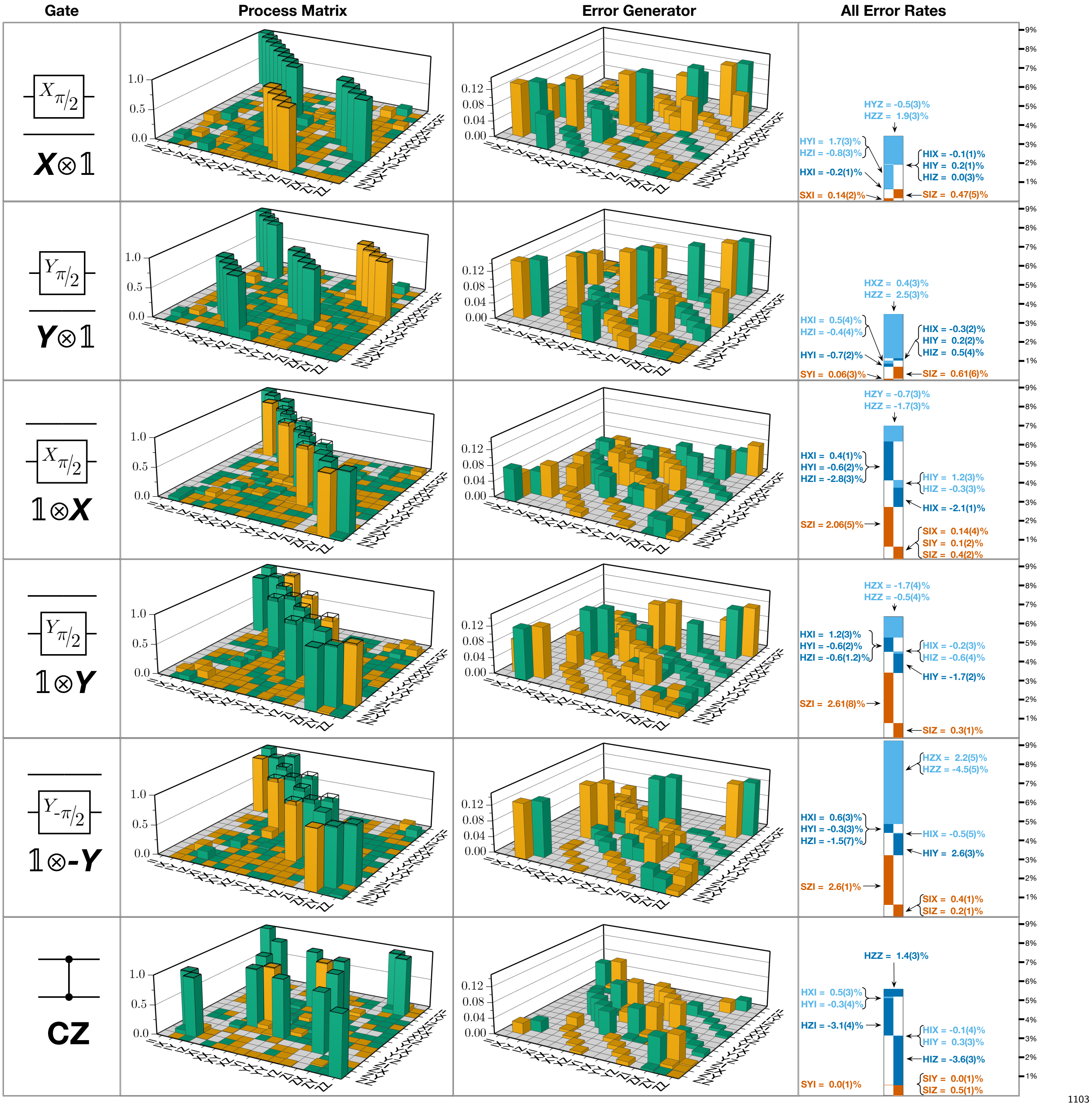
Extended Data Fig. 5 | Nuclear spin quantum jumps caused by ionization shock. The electron and nuclear spin readout relies upon spin-dependent charge tunnelling between the donors and the SET island. If the electron tunnels out of the two-donor system, the hyperfine interactions A_1, A_2 suddenly drop to zero. If A_1 and A_2 include an anisotropic component (e.g. due to the non-spherical shape of the electron wavefunction which results in nonzero dipolar fields at the nuclei), the ionisation is accompanied by a sudden change in the nuclear spin quantisation axes ("ionisation shock"), and can result in a flip of the nuclear spin state. We measure the nuclear spin flips caused by ionisation shock by forcibly loading and unloading an electron from the 2P cluster every 0.8 ms. **a,** For qubit 1 with $A_1 = 95$ MHz, the flip rate is $\Gamma_1 = 2.8 \times 10^{-6} \frac{\mathrm{N_{flip}}}{\mathrm{N_{ion}}}$. **b,** For qubit 2 with $A_2 = 9$ MHz, the flip rate is $\Gamma_2 = 4.0 \times 10^{-7} \frac{\mathrm{N_{flip}}}{\mathrm{N_{ion}}}$. This means that the nuclear spin readout via the electron ancilla is almost exactly quantum non-demolition. From this data, we also extract an average time between random nuclear spin flips of 283 seconds for qubit 1, and 2000 seconds for qubit 2. The extremely low values of Γ – comparable to those observed in single-donor systems – are the reason why we can reliably operate the two ^{31}P nuclei as high-fidelity qubits.



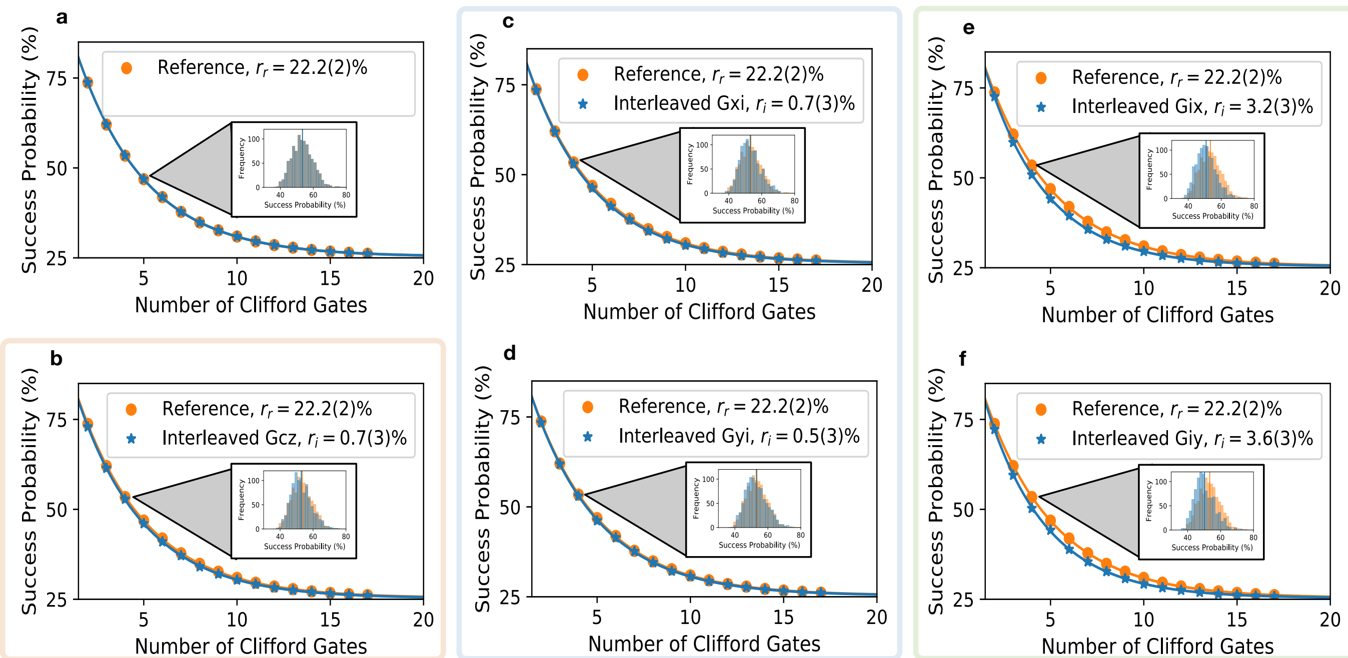
Extended Data Fig. 6 | CNOT and zero-CNOT nuclear two-qubit gates. We perform Rabi oscillation on the control qubit followed by the application of **a,** zCNOT or **b,** CNOT gates. The two qubits are initialized in the $|\Downarrow\Downarrow\rangle \equiv |11\rangle$ state. We observe the Rabi oscillations of both qubits in phase for zCNOT and out of phase for CNOT. At every odd multiple of $\pi/2$ rotation of the control qubit the Bell states are created.



Extended Data Fig. 7 | Two-qubit gate set tomography. a, Measurement circuit for the two-qubit gate set tomography. A modified version of this circuit has been used for Bell state tomography. The green box prepares the qubit 2 in the $|\Uparrow\rangle$ state, then the orange box prepares the qubit 1 in the $|\Uparrow\rangle$ state. The readout step in the blue box (see Methods) determines whether the $|\Uparrow\Uparrow\rangle$ state initialization was successful. Only then the record will be saved. The electron spin is prepared in $|\downarrow\rangle$ during the nuclear spin readout process. Subsequently, the GST sequence is executed. The red box indicates the Q1,Q2 readout step. The total duration of the pulse sequence is 120 ms, of which nuclear spin initialization is 8.6 ms (green and yellow), initial nuclear spin readout is 26.5 ms (blue), 3 ms delay is added for electron initialization (between blue and purple), GST circuit is 10 μs - 300 μs (purple), and nuclear readout is 80 ms (orange). **b,** Measurement results for individual two-qubit gate set tomography circuit. The first 145 circuits estimate the preparation and measurement fiducials, and the subsequent circuits are ordered by increasing circuit depth. At the end of a circuit, there are three situations for the target state populations: 1) the population is entirely in one state, while all others are zero; 2) the population is equally spread over two states, while the other two are zero; 3) the population is equally spread over all four states. The measured state populations for the different circuits therefore congregate around the four bands 0, 0.25, 0.5, and 1, as indicated by black dashed lines.


Gate
Process Matrix
Error Generator
All Error Rates
HYZ = -0.5(3)%
HZZ = 1.9(3)%
HYI = 1.7(3)%
HZI = -0.8(3)%
HXI = -0.2(1)%
SXI = 0.14(2)%
HIX = -0.1(1)%
HIY = 0.2(1)%
HIZ = 0.0(3)%
SIZ = 0.47(5)%
HXZ = 0.4(3)%
HZZ = 2.5(3)%
HXI = 0.5(4)%
HZI = -0.4(4)%
HYI = -0.7(2)%
SYI = 0.06(3)%
HIX = -0.3(2)%
HIY = 0.2(2)%
HIZ = 0.5(4)%
SIZ = 0.61(6)%
HZY = -0.7(3)%
HZZ = -1.7(3)%
HXI = 0.4(1)%
HYI = -0.6(2)%
HZI = -2.8(3)%
HIY = 1.2(3)%
HIZ = -0.3(3)%
HIX = -2.1(1)%
SZI = 2.06(5)%
SIX = 0.14(4)%
SIY = 0.1(2)%
SIZ = 0.4(2)%
HZX = -1.7(4)%
HZZ = -0.5(4)%
HXI = 1.2(3)%
HYI = -0.6(2)%
HZI = -0.6(1.2)%
HIX = -0.2(3)%
HIZ = -0.6(4)%
HIY = -1.7(2)%
SZI = 2.61(8)%
SIZ = 0.3(1)%
HZX = 2.2(5)%
HZZ = -4.5(5)%
HXI = 0.6(3)%
HYI = -0.3(3)%
HZI = -1.5(7)%
HIX = -0.5(5)%
HIY = 2.6(3)%
SZI = 2.6(1)%
SIX = 0.4(1)%
SIZ = 0.2(1)%
HZZ = 1.4(3)%
HXI = 0.5(3)%
HYI = -0.3(4)%
HZI = -3.1(4)%
HIX = -0.1(4)%
HIY = 0.3(3)%
HIZ = -3.6(3)%
SYI = 0.0(1)%
SIY = 0.0(1)%
SIZ = 0.5(1)%
CZ
1103

Extended Data Fig. 8 | Estimated gate set, from process matrices to error rates. Experimental GST data were analyzed using pyGSTi to obtain self-consistent maximum likelihood estimates of 2-qubit process matrices for all 6 elementary gates. These are represented ("Process Matrix" column) in a gauge that minimizes their average total error, as superoperators in the 2-qubit Pauli basis. Green columns indicate positive matrix elements, orange ones are negative. Wireframe sections indicate differences between estimated and ideal (target) process matrices. Those process matrices can be transformed to error generators ("Error Generator" column) that isolate those differences, and are zero if the estimated gate equals its target. Each gate's error generator was decomposed into a sparse sum of Hamiltonian and stochastic elementary error generators [33]. Those rates are depicted ("All Error Rates" column) as contributions to the gate's total error, with one-sigma uncertainties indicated in parentheses. Each non-vanishing elementary error rate (error generators are denoted "H" or "S" followed by a Pauli operator) is listed, and identified with its role in the total error budget (reproduced from Figure 3). Orange bars indicate stochastic errors, dark blue indicate coherent errors that are intrinsic to the gate, and light blue indicate relational coherent errors that were assigned to this gate. Total height of the blue region indicates the total coherent error, but because coherent error amplitudes add in quadrature, individual components' heights are proportional to their quadrature.



Extended Data Fig. 9 | Simulation of standard and interleaved randomized benchmarking (RB). All simulated RB experiments used 2-qubit Clifford subroutines compiled from the 6 native gates, requiring (on average) 14.58 individual gate operations per 2-qubit Clifford. **a**, Standard randomized benchmarking, simulated using the GST-estimated gate set, yields a "reference" decay rate of $r_r = 22.2(2)\%$, suggesting an average per-gate error rate of $r_r/14.58 \approx 1.5\%$. One-sigma confidence intervals are indicated in parentheses. **b-f**, Simulated interleaved randomized benchmarking for the CZ gate, and 1-qubit $X_{\pi/2}$ and $Y_{\pi/2}$ gates on each qubit, yielded interleaved decay rates $r_r + r_i$. For each experiment, 1000 random Clifford sequences were generated, at each of 15 circuit depths m, and simulated using the GST process matrices. Exact probabilities (effectively infinitely many shots of each sequence) were recorded. Inset histograms show the distribution over 1000 random circuits at m=4. Observed decays are consistent with each gate's GST-estimated infidelities – e.g. $1 - F = 0.79\%$ for the C-Z gate (b). Performing these exact RB experiments in the lab would have required running 90000 circuits to estimate a single parameter (r_i) for each gate to the given precision of $\pm 0.25\%$. Using fewer (< 1000) random circuits at each m would yield lower precision. GST required only 1500 circuits to estimate *all* error rates to the same precision.

	$\lvert\Uparrow\Uparrow\rangle$	$\lvert\Uparrow\Downarrow\rangle$	$\lvert\Downarrow\Uparrow\rangle$	$\lvert\Downarrow\Downarrow\rangle$
Pr($\lvert\Uparrow\Uparrow\rangle$)	99.75(3)%	0.53(7)%	0.53(6)%	0.52(4)%
Pr($\lvert\Uparrow\Downarrow\rangle$)	0.04(1)%	99.09(8)%	0.02(1)%	0.06(2)%
Pr($\lvert\Downarrow\Uparrow\rangle$)	0.20(3)%	0.18(3)%	97.73(10)%	0.20(5)%
Pr($\lvert\Downarrow\Downarrow\rangle$)	0.02(1)%	0.20(3)%	1.72(8)%	99.22(6)%

Extended Data Fig. 10 | Estimated state preparation and measurement (SPAM) error rates. In the GST analysis, the system's initial state was represented by a 4×4 density matrix ρ, and the final measurement/readout by a 4-element 4×4 POVM (positive operator-valued measure) $\{E_{\Uparrow\Uparrow}, E_{\Uparrow\Downarrow}, E_{\Downarrow\Uparrow}, E_{\Downarrow\Downarrow}\}$ with $E_j \geq 0$ and $\sum_j E_j = I$. We quantified the overall quality of the SPAM operations by using the GST estimate to compute the table of conditional probabilities shown here. Each cell shows the estimated probability of a particular readout (e.g. $\Uparrow\Uparrow$) given (imperfect) initialization in a particular state (e.g. $\lvert\Downarrow\Downarrow\rangle$). The $\lvert\Uparrow\Uparrow\rangle$ column can be read out directly from the estimate, since the experiment initalized into $\lvert\Uparrow\Uparrow\rangle$. Other states must be prepared by applying $X_{\pi/2}$ or $Y_{\pi/2}$ pulses. These add additional error, which should not be attributed to SPAM operations. To correct for this, we simulated ideal unitary rotation of the real $\lvert\Uparrow\Uparrow\rangle$ state into each of the other 3 states by (1) taking the GST-estimated $X_{\pi/2}$ gates on each qubit and removing all intrinsic errors from them, and (2) simulating a circuit comprising initialization in ρ, an appropriate sequence of those idealized gates, and readout according to $\{E_j\}$. The resulting analysis shows probabilities of all but one readout error to be below 1%, which is unprecedented in semiconductor spin qubit systems, and competitive with the state of the art in other physical platforms.