

Robust multi-qubit quantum network node with integrated error detection

Authors:

P.-J. Stas^{1*}, Y. Q. Huan^{1*}, B. Machielse^{1,2*}, E. N. Knall³, A. Suleymanzade¹, B. Pingault^{3,4,5}, M. Sutula¹, S. W. Ding³, C. M. Knaut¹, D. R. Assumpcao³, Y.-C. Wei¹, M. K. Bhaskar^{1,2}, R. Riedinger^{1,6,7}, D. D. Sukachev^{1,2}, H. Park^{1,8}, M. Lončar³, D. S. Levonian^{1,2}, M. D. Lukin^{1†}

Affiliations:

¹Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

²AWS Center for Quantum Networking, Boston, Massachusetts 02210, USA

³John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

⁴QuTech, Delft University of Technology, 2600 GA Delft, The Netherlands

⁵Kavli Institute of Nanoscience Delft, Delft University of Technology, 2600 GA Delft, The Netherlands

⁶Institut für Laserphysik und Zentrum für Optische Quantentechnologien, Universität Hamburg, 22761 Hamburg, Germany

⁷The Hamburg Centre for Ultrafast Imaging, 22761 Hamburg, Germany

⁸Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA

*These authors contributed equally to this work.

†To whom correspondence should be addressed; E-mail: lukin@physics.harvard.edu.

Abstract: Long-distance quantum communication and networking require quantum memory nodes with efficient optical interfaces and long memory times. We report the realization of an integrated two-qubit network node

based on silicon-vacancy centers (SiVs) in diamond nanophotonic cavities. Our qubit register consists of the SiV electron spin acting as a communication qubit and the strongly coupled ^{29}Si nuclear spin acting as a memory qubit with a quantum memory time exceeding two seconds. By using a highly strained SiV with suppressed electron spin-phonon interactions, we realize electron-photon entangling gates at elevated temperatures up to 1.5 K and nucleus-photon entangling gates up to 4.3 K. Finally, we demonstrate efficient error detection in nuclear spin-photon gates by using the electron spin as a flag qubit, making this platform a promising candidate for scalable quantum repeaters.

One Sentence Summary: A quantum network node is realized combining a long-lived memory, robust optical interface and integrated error detection.

Main Text: The ability to distribute quantum information over extended distances (1, 2) constitutes an important enabling technology in quantum information science, with applications in quantum key distribution (3,4), nonlocal sensing (5), and distributed quantum computation (6, 7). A key requirement for the realization of such long-distance quantum networking involves the development of quantum repeaters (8,9) to mitigate photonic qubit loss during the transmission over extended distances. Proposed repeater architectures require nodes containing multiple qubits that can collect, store, and process information communicated via photonic channels (10).

Color centers in diamond nanophotonic structures (11, 12, 13, 14) have recently emerged as promising candidates for realizing such nodes due to their long coherence times, high-fidelity single-qubit gates, and efficient photonic interfaces. The integration of all these features into a single device has led to the demonstration of memory-enhanced quantum communication (15). However, scalable implementation of quantum repeaters using this approach requires deterministic access to additional long-lived quantum memory qubits to perform entanglement swap-

ping, purification and error detection (8, 10). Furthermore, the technical complexity and costs associated with the need for operation at dilution refrigerator temperatures to avoid phonon-induced dephasing (16, 17) has stimulated efforts to identify quantum memory systems capable of operating within a wider range of environmental conditions (18, 19).

In this Report, we address these challenges by using a highly strained silicon-vacancy (SiV) color center (16, 20) featuring the ^{29}Si isotope, with its nuclear spin (21) serving as a deterministic long-lived memory qubit. We demonstrate full control of this two-qubit register with a ^{29}Si quantum memory storage time of over 2 seconds. Selective qubit readout via a novel phase-based readout method enables multiple electron state resets before deterioration of the nuclear spin memory. Using this system, we achieve electron spin-photon entanglement at temperatures up to 1.5 K and direct nuclear-photon entanglement up to 4 K. Finally, we combine these elements to demonstrate nuclear-photon spin entanglement with integrated error detection using the electronic spin as a flag qubit to detect operational errors.

The ^{29}SiV in an external magnetic field constitutes a two-qubit system of four spin states with nondegenerate transition frequencies (Fig. 1A). The electron and nuclear spin qubits are coherently controlled using microwave (MW) and radiofrequency (RF) pulses, respectively, delivered via gold coplanar waveguides (Fig. 1B). The SiV is embedded in a nanophotonic cavity (Fig. 1C) enhancing optical transitions to the excited state manifold at 737 nm (Fig. 1A, wavy arrows) used for state readout and spin-photon entanglement. The SiV-cavity system exhibits high-contrast spin-dependent reflection spectra (Fig. 1D) enabled by the strong cavity coupling (cooperativity $C = 1.6$ (22)).

High-fidelity resonant readout of the electron state is achieved by measuring the reflected intensity of a laser at the frequency of maximum reflection intensity contrast (black line in Fig. 1D) (11, 23). However, the nuclear qubit experiences dephasing from the laser during readout, with the highest decoherence rate occurring when the laser is near the resonant readout

frequency (22). To enable selective readout of the electronic spin qubit, we probe the system at laser frequencies where minimal nuclear dephasing occurs and make use of the electron spin-dependent phase – instead of intensity – of reflected photons for readout. Using an electro-optic modulator (EOM) to generate sidebands (Fig. 1E), we send two tones (blue and magenta lines in Fig. 1D) along the same path and measure the spin-dependent reflected phase difference via the resulting beating pattern to determine the electron state (Fig. 1F) (22). With this phase-based readout, the electron can be read 14 times at 95% fidelity before causing a $1/e$ loss of nuclear coherence. This constitutes an 8-fold improvement over the resonant readout method (Fig. 1G), with further improvements expected through cavity design and magnetic field optimization (22).

The two-qubit ^{29}SiV system is fully controlled by selectively driving the four single-spin-flipping transitions (Fig. 1A) to implement the four possible controlled-NOT (CNOT) gates: two electron-flipping gates $C_n\text{NOT}_e$ (MW1) and $\overline{C_n\text{NOT}_e}$ (MW2), and two nucleus-flipping gates $C_e\text{NOT}_n$ (RF1) and $\overline{C_e\text{NOT}_n}$ (RF2), where the absence (presence) of the overbar indicates conditioning on the control spin being in the down (up) state. We measure a fidelity of $99.9 \pm 0.1\%$ for the $C_n\text{NOT}_e$ gate with a gate time of 30.0 ns. As the time to drive a nuclear π rotation ($23 \mu\text{s}$) is longer than the electron dephasing time $T_{2,e}^*$ ($5 \mu\text{s}$) (22), transfer of electron superposition states onto the nucleus is not possible with direct driving for the $C_e\text{NOT}_n$ gate. To circumvent this problem, we increase the electron coherence time using a dynamical decoupling sequence and interleave it with step-wise nuclear rotations (24, 25) to implement a decoupled $C_e\text{NOT}_n$ (Fig. 2A). Specifically, we apply RF1 in the first window, and alternate it with applying RF2 after every subsequent unconditional electron π pulse to account for the flipping electron state while keeping the nuclear rotation conditional on the same initial electron state, achieving a gate fidelity of $93.7 \pm 0.7\%$ (22) within a gate time of $29.2 \mu\text{s}$. With the decoupled $C_e\text{NOT}_n$ we can swap a superposition state from the electron onto the nuclear spin state (Fig. 2B), where it can be stored using an XY8 decoupling sequence. We find that the

nuclear coherence time scales with n , the number of decoupling pulses applied, as $T_{2,n} \propto n^{0.53}$ with a maximum measured nuclear memory time of $T_{2,n} = 2.1 \pm 0.1$ seconds (Fig. 2C) when using a 128-XY8 sequence.

To enable robust spin-photon entanglement, we use a SiV with large residual strain (ground state splitting $\Delta_{GS} = 554$ GHz (22)), which greatly suppresses the rate of thermal decoherence processes (20), enabling operation at 1.5 K without a significant reduction in $T_{2,e}$ (Fig. 3A). We implement high-temperature spin-photon entangling gates between the electron spin and time-bin qubits in the $\{|e\rangle, |l\rangle\}$ basis, corresponding to the presence of a photon in either the early or late time-bin. The entangling gate sequence (11, 15) shown in Fig. 3B generates the entangled state $(|e \downarrow_e\rangle + |l \uparrow_e\rangle)/\sqrt{2}$ conditioned on the detection of a single reflected photon. We find a Bell state fidelity of $F_{e\gamma} = 0.91 \pm 0.02$ (0.90 ± 0.01) at 0.1 K (1.5 K) (Fig. 3C), which is primarily limited by residual reflections from the $|\downarrow_e\rangle$ state and imperfect photonic state measurements.

To further extend the entanglement capabilities of our spin-photon interface, we introduce a novel PHOton-Nucleus Entangling (PHONE) gate that directly entangles the ^{29}Si nuclear spin with a photonic qubit using only fast MW gates and allows operating temperatures up to 4.3 K. In this scheme, the photon-nucleus-electron system is initialized in the $(|e\rangle + |l\rangle)(|\uparrow_n\rangle + |\downarrow_n\rangle)|\downarrow_e\rangle/2$ state and we apply a sequence of CNOTs as shown in Fig. 3D such that the SiV reflectivity in each time bin is conditional on the nuclear state; the final entangled state is given by $(|e \uparrow_n\rangle + |l \downarrow_n\rangle)|\downarrow_e\rangle/\sqrt{2}$. We measure the resulting Bell state fidelity to be $F_{n\gamma} = 0.85 \pm 0.02$ (0.66 ± 0.02) at 0.1 K (4.3 K) (Fig. 3E, F), with dominant infidelities due to MW gate errors and the nuclear state-dependence of the optical transition frequency which prevents the photon frequency from being tuned to maximize contrast for both nuclear spin states simultaneously (Fig. 1D) (22).

A key feature of the PHONE gate is that the electron should always remain in the $|\downarrow_e\rangle$

state after a successful gate application. Since the electron spin mediates the interface between the photon and the nucleus, electron spin flips can be used as an integrated error witness to detect gate errors. Similar to flag qubits in error correction protocols (26), PHONE gate errors can be reduced at the cost of some gate failure probability (27) by measuring the state of the electron spin qubit (Fig. 4A). Specifically, if the electron spin is measured to be in the $|\uparrow_e\rangle$ state following a PHONE gate application, the prepared photon-nuclear spin entangled state is discarded. As shown in Fig. 4B, the use of the procedure results in a Bell state fidelity increase of 2% (7%) to $F_{n\gamma} = 0.87 \pm 0.02$ (0.71 ± 0.02) with an error detection rate of 8.4% (13.9%) at 0.1 K (4.3 K). In Fig. 4C we combine all these components to implement a PHONE gate at 4.3 K with error detection and store the spin-photon entangled state in the nuclear spin memory using an echo sequence (Fig. 4A) for over 2.5 ms above the threshold value of 50%, as compared to 1.5 ms without error detection. The error detection protocol detects MW gate errors and $T_{1,e}$ and $T_{2,e}$ -limited depolarization and dephasing, and the large gain in fidelity at 4.3 K is due to the greater contribution of the detectable errors from short $T_{1,e}$ and $T_{2,e}$ times at this temperature (22). Conversely, the improvement from error detection at 0.1 K is limited as the infidelity is dominated by non-detectable errors from photon state measurement and SiV optical contrast.

These observations open up several new avenues for realizing quantum networks and exploring their applications. The access to an additional memory qubit directly enables improved memory-enhanced quantum key distribution (15) by extending the time window for photons arriving from different communicating parties to the memory node for asynchronous Bell measurement, whereas the phase readout protocol facilitates electron resets between entanglement attempts while information is stored on the nucleus. A nuclear memory can additionally expand the capabilities of the SiV as a single photon source for the creation of photonic cluster states (28,29). The methods demonstrated here can also enable the deployment of scalable SiV-

based quantum repeater networks. Based on a large-scale survey of a second chip fabricated using the current process (22), we find that more than 11% of all SiVs are sufficiently strained for operation at 1.5 K, indicating the possibility for a significant number of highly-strained devices in nanophotonic cavities capable of operating at elevated temperatures. Moreover, recently demonstrated techniques for strain tuning of nanocavities (30) can enable fully deterministic access to high-strain operation, eliminating the need for dilution refrigerators. The fidelity gain from our error detection can be more significant for more complex spin-photon entangling sequences, such as PHONE-type gates entangling successive photons with the nucleus or entangling a photon with multiple nuclei strongly coupled to the SiV. It also opens opportunities to operate more effectively in regimes where the SiV electron coherence properties deteriorate, including at higher temperatures as demonstrated here, and in a misaligned magnetic field as is required for acoustic spin control (31) and single photon generation (29). The use of cavities with higher cooperativity as demonstrated previously (15) should allow for higher fidelity and efficiency of electron-spin photon and PHONE gates, as well as improved ^{29}Si state preservation during electron readout (22). Finally, nearby ^{13}C spins, such as one associated with the SiV used in this work (22), can be used as additional memory resources (24, 32). Apart from realizing multi-node quantum network protocols (8, 9, 10), these systems can also allow for the generation of complex photonic tree cluster states enabling robust one-way long-distance quantum communication (33).

References

1. H. J. Kimble, *Nature* **453**, 1023 (2008).
2. S. Wehner, D. Elkouss, R. Hanson, *Science* **362** (2018).
3. A. K. Ekert, *Physical Review Letters* **67**, 661 (1991).

4. H.-K. Lo, M. Curty, K. Tamaki, *Nature Photonics* **8**, 595 (2014).
5. E. Khabiboulline, J. Borregaard, K. D. Greve, M. Lukin, *Physical Review Letters* **123**, 070504 (2019).
6. H. Buhrman, H. Röhrig, *Mathematical Foundations of Computer Science 2003* (Springer Berlin Heidelberg, 2003), pp. 1–20.
7. D. Cuomo, M. Caleffi, A. S. Cacciapuoti, *IET Quantum Communication* **1**, 3 (2020).
8. W. Dür, H.-J. Briegel, J. I. Cirac, P. Zoller, *Physical Review A* **59**, 169 (1999).
9. L. Jiang, *et al.*, *Physical Review A* **79**, 032325 (2009).
10. L. Childress, J. M. Taylor, A. S. Sørensen, M. D. Lukin, *Physical Review Letters* **96**, 070504 (2006).
11. C. T. Nguyen, *et al.*, *Physical Review B* **100**, 165428 (2019).
12. E. R. Schmidgall, *et al.*, *Nano Letters* **18**, 1175 (2018).
13. A. E. Rugar, *et al.*, *ACS Photonics* **9**, 2356–2361 (2020).
14. X. Guo, *et al.*, *Nano Letters* **21**, 10392 (2021).
15. M. K. Bhaskar, *et al.*, *Nature* **580**, 60 (2020).
16. K. D. Jahnke, *et al.*, *New Journal of Physics* **17**, 043011 (2015).
17. J. N. Becker, *et al.*, *Physical Review Letters* **120**, 1 (2018).
18. B. C. Rose, *et al.*, *Science* **361**, 60 (2018).
19. M. E. Trusheim, *et al.*, *Physical Review Letters* **124**, 023602 (2020).

20. S. Meesala, *et al.*, *Physical Review B* **97**, 205444 (2018).
21. B. Pingault, *et al.*, *Nature Communications* **8**, 1 (2017).
22. See Supplemental Material at [URL inserted by publisher] for further details on theoretical analysis and additional supporting experimental results.
23. D. S. Levonian, *et al.*, *Phys. Rev. Lett.* **128**, 213602 (2022).
24. C. Bradley, *et al.*, *Physical Review X* **9**, 031045 (2019).
25. T. van der Sar, *et al.*, *Nature* **484**, 82 (2012).
26. R. Chao, B. W. Reichardt, *Physical Review Letters* **121**, 050502 (2018).
27. J. Borregaard, P. Kómár, E. Kessler, A. Sørensen, M. Lukin, *Physical Review Letters* **114**, 110502 (2015).
28. N. H. Lindner, T. Rudolph, *Physical Review Letters* **103**, 113602 (2009).
29. E. N. Knall, *et al.*, *arXiv:2201.02731* (2022).
30. B. Machielse, *et al.*, *Physical Review X* **9**, 031022 (2019).
31. S. Maity, *et al.*, *Nature Communications* **11**, 193 (2020).
32. N. Kalb, *et al.*, *Science* **356**, 928 (2017).
33. J. Borregaard, *et al.*, *Physical Review X* **10**, 021071 (2020).
34. M. J. Burek, *et al.*, *Physical Review Applied* **8**, 024026 (2017).
35. C. Chia, B. Machielse, A. Shams-Ansari, M. Lončar, *Opt. Express* **30**, 14189 (2022).

36. C. M. Knaut, *et al.*, pylabnet - Client-server, python-based laboratory software, <https://github.com/lukingroup/pylabnet> (2021).
37. A. Reiserer, G. Rempe, *Reviews of Modern Physics* **87**, 1379 (2015).
38. B. Smeltzer, L. Childress, A. Gali, *New Journal of Physics* **13** (2011).
39. D. Sukachev, *et al.*, *Physical Review Letters* **119**, 223602 (2017).
40. R. Orbach, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **264**, 458 (1961).
41. T. Müller, *et al.*, *Nature Communications* **5**, 3328 (2014).
42. C. Hepp, thesis, Universität des Saarlandes (2014).
43. L. A. O'Neill, B. Joecker, A. D. Baczewski, A. Morello, *Applied Physics Letters* **119**, 174001 (2021).
44. Y. Sagi, R. Pugatch, I. Almog, N. Davidson, *Physical Review Letters* **104**, 253003 (2010).
45. P. W. Anderson, *Journal of the Physical Society of Japan* **9**, 316 (1954).
46. L. Jiang, *et al.*, *Physical Review Letters* (2008).
47. N. Kalb, P. C. Humphreys, J. J. Slim, R. Hanson, *Physical Review A* **97**, 062330 (2018).
48. C. A. Fuchs, thesis, University of New Mexico (1996).

Acknowledgments: We thank J. Borregaard for discussions, and J. MacArthur for assistance with electronics. This work was supported by the NSF, CUA, DoD/ARO DURIP, AFOSR MURI, and DOE (award no. DE-SC0020115). Devices were fabricated at Harvard CNS, NSF

award no. 1541959. Y.Q.H acknowledges support from the A*STAR National Science Scholarship. D.A., E.N.K., and B.M. acknowledge support from an NSF GRFP No. DGE1745303. B.P. acknowledges support through a Marie Skłodowska-Curie fellowship from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 840968 (COHESiV). M.S. acknowledges funding from the NASA Space Technology Graduate Research Fellowship Program. R. R. acknowledges support from the Alexander von Humboldt Foundation and the Cluster of Excellence 'Advanced Imaging of Matter' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994.

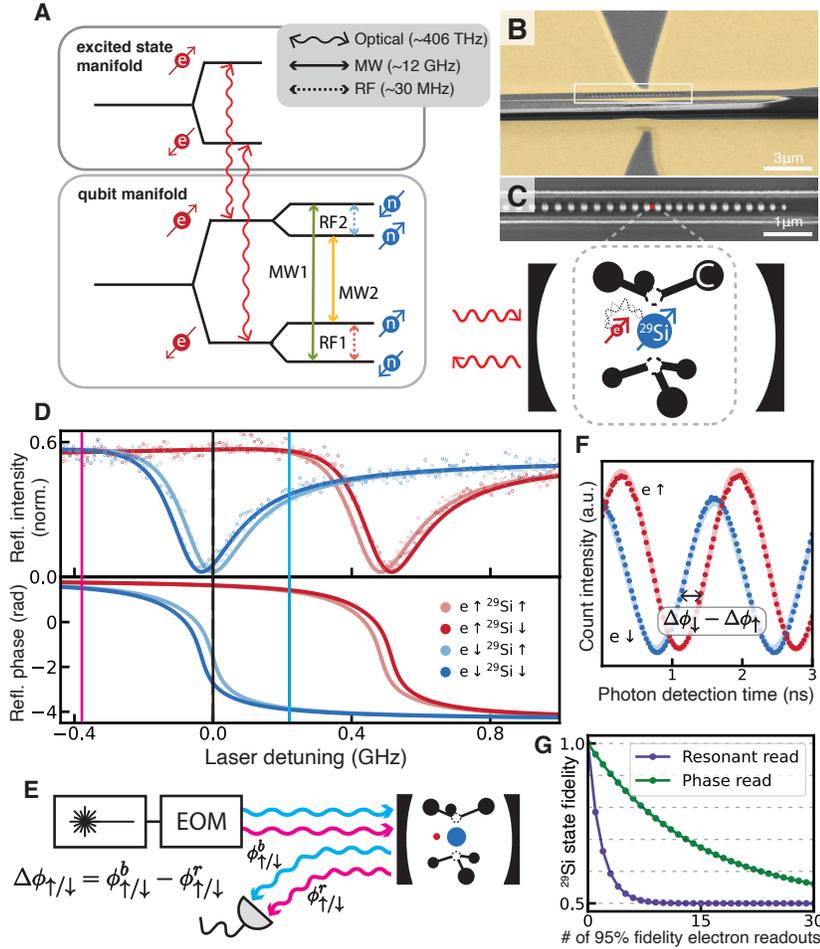


Fig. 1. Quantum network node based on ^{29}SiV . (A) Energy level diagram of the ^{29}SiV . Allowed magnetic dipole transitions are shown by arrows in the qubit manifold for the case when the **B**-field is oriented along the symmetry axis. (B) False color scanning electron microscope (SEM) image of the device, showing the gold coplanar waveguides in yellow. (C) Zoomed in SEM image of the nanophotonic cavity in which the ^{29}SiV is located. (D) Spin-dependent reflection intensity (top) and phase (bottom) as a function of laser frequency. (E) Experimental setup for phase readout of the electron spin state. (F) Measured spin-dependent phase shift of the beating pattern between sidebands. (G) Coherence of a nuclear superposition state as a function of the number of electron readouts for resonant and phase readouts.

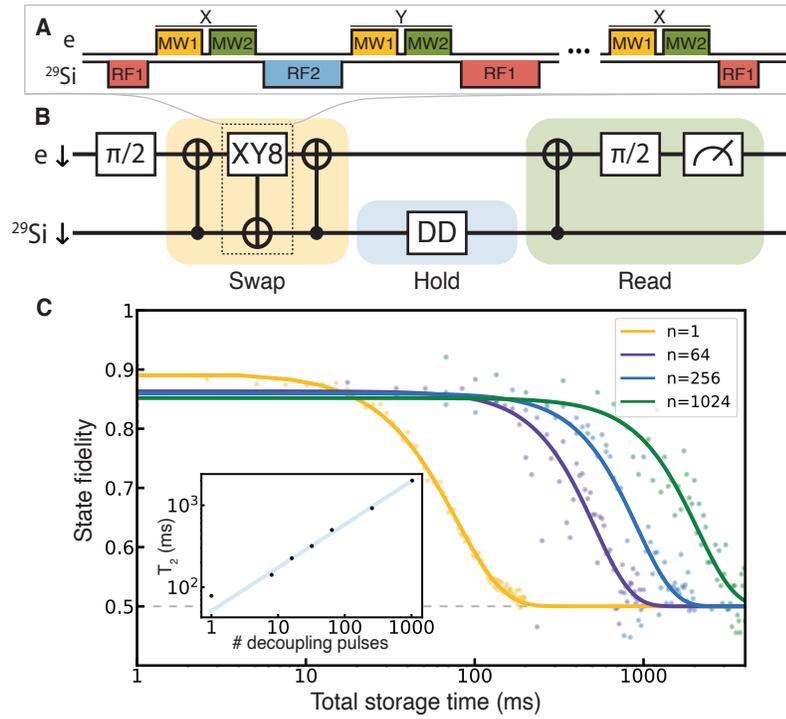


Fig. 2. Long-lived quantum memory based on ^{29}Si nuclear spin. (A) Pulse sequence for a decoupled $C_e\text{NOT}_n$ gate. (B) Sequence for swapping the electron state onto the nucleus, after which a nuclear XY8 decoupling sequence with $n = 1, 64, 256, 1024$ pulses is applied and the nuclear state is measured through resonant readout of the electron. (C) Measured nuclear state fidelity after sequence (B). Inset: $T_{2,n}$ as a function of the number of decoupling pulses n fitted to $T_{2,n} \propto n^\alpha$ where $\alpha = 0.53$.

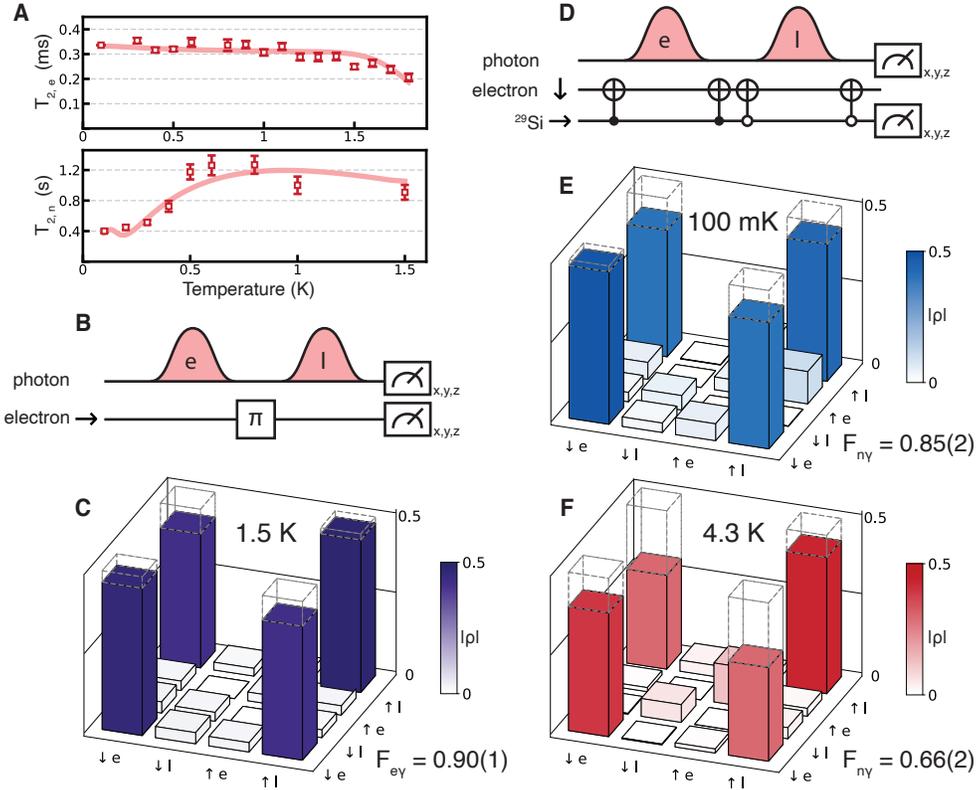


Fig. 3. Spin-photon entanglement at elevated temperatures. (A) Coherence time of the electron (top) and nucleus (bottom) obtained using 8-XY8 sequences as a function of temperature. Initial increase of the nuclear $T_{2,n}$ is due to motional averaging of the noise bath. Solid lines are fits to decoherence models of the spin environment described in (22). (B) Electron spin-photon gate implementation with (C) reconstructed density matrix at 1.5 K using resonant electron readout. (D) Circuit diagram for the PHOTon-Nucleus Entangling (PHONE) gate. CNOT gates with black and white control dots refer to $C_n\text{NOT}_e$ and $C_n\text{NOT}_e$ gates respectively. Reconstructed density matrices for photon-nuclear spin entanglement at (E) 100 mK and (F) 4.3 K.

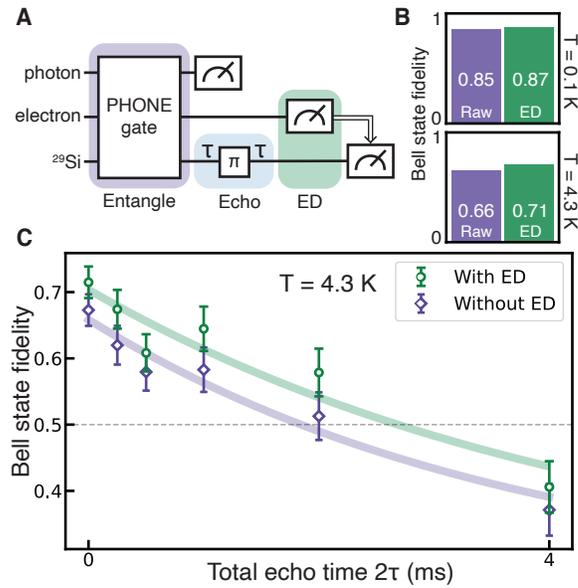


Fig. 4. Spin-photon entanglement with integrated error detection. (A) Gate sequence for entanglement between a photon and nucleus and subsequent state storage with error detection (ED) based on resonant readout of the electronic spin. (B) Bell state fidelity improvement with ED at 0.1 K and 4.3 K. (C) Bell state fidelity as a function of storage time for the sequence shown in (A) before and after error detection. Fits are exponential curves decaying to the maximally mixed state fidelity $\langle \Phi^+ | \rho_{\text{mm}} | \Phi^+ \rangle = 0.25$ as the nuclear lifetime $T_{1,n}$ is comparable to the coherence time; fitted decay constants are 4.5 (3.8) ms with (without) ED.

Supplementary materials

Materials and Methods

Supplementary Text

Figs. S1 to S15

Tables S1 to S5

References (34 – 48)

Supplemental Materials for **Robust multi-qubit quantum network node with integrated error detection**

1 Sample Fabrication

The nanophotonic crystal cavities utilized in this manuscript were patterned out of bulk diamond utilizing techniques previously reported in (11). Subsequent implantation of ^{29}Si through apertures patterned into a resist mask, followed by high temperature annealing, enables incorporation of color centers into these photonic crystals with a high degree of spatial accuracy. The optical properties of these devices can be measured through a hydrofluoric acid wet-etched tapered optical fiber (34) brought into contact with the tapered end of the diamond devices. By measuring the reflection spectra of the photonic crystal devices inside a dilution refrigerator operating at 100 mK, strong modulation of the cavity properties by the presence of these color centers can be observed, indicating relatively high cooperativity ($C = 4g^2/\kappa\gamma = 1.6$; see section on cavity-QED model for parameter values) coupling of the emitters with photons in the cavity mode. Further improvements of cooperativities can be achieved through optimization of the etch mask (35). Candidate devices can then be identified and, after removal from the cryostat, deposited with coplanar waveguides which are patterned using electron beam lithography (Fig. 1B).

2 Experimental setup

We perform all measurements in a dilution refrigerator (BlueFors BF-LD250) with an effective base temperature of 100 mK. The dilution refrigerator is equipped with a superconducting vector magnet (American Magnets Inc. 6-1-1 T), a home-built free-space wide-field microscope with an objective lens, piezo positioners (Attocube ANPx101 and ANPx311 series), and fiber and microwave feedthroughs (gold coplanar waveguides). A gas line with attached heaters is used for tuning of the nanocavity resonance through nitrogen gas condensation.

The optical path consists of two separate 737 nm Ti:Sapphire lasers (TiSaph, M Squared SolsTiS). One of these lasers is used exclusively for phase readout and is tuned to the phase readout carrier frequency, while the other is tuned to the resonant readout frequency and is used for resonant readout, photonic qubit generation, and time delay interferometer (TDI) locking. A separate green diode laser (532 nm, Thorlabs) is used to reset the SiV electron charge state when it ionizes. The acousto-optic modulators (AOMs) on the optical path are Gooch & Housego R1520 and the EOMs are EOSpace AZ-OK5-10-PFA-PFA-637 Mach-Zehnder amplitude modulators. All of the laser paths (with the exception of the TDI locking beam) are combined into a single-mode S630-HP fiber (Thorlabs) and sent into the dilution fridge where

it couples light adiabatically in and out of the diamond nanocavity via a tapered fiber (Thorlabs 630-HP). Reflected photons from the device travel along the same path through a 99:1 beam splitter where it is collected either on an SNSPD (Photon Spot) or a pair of avalanche photodiodes (APDs, Excelitas SPCM-780-14-FC). The APDs are used for all experiments at 4 K and the spin-photon entangling gates; the SNSPDs are used for all other experiments. Counts from these photon detectors are recorded on a time-tagger (Swabian Instruments Time Tagger Ultra) as well as an integrated counter on the arbitrary waveform generator (AWG) for state readout and conditional logic. A pair of photodiodes are placed along the optical path to calibrate the coupling efficiency of the fiber with the device by measuring the ratio of incoming to reflected light.

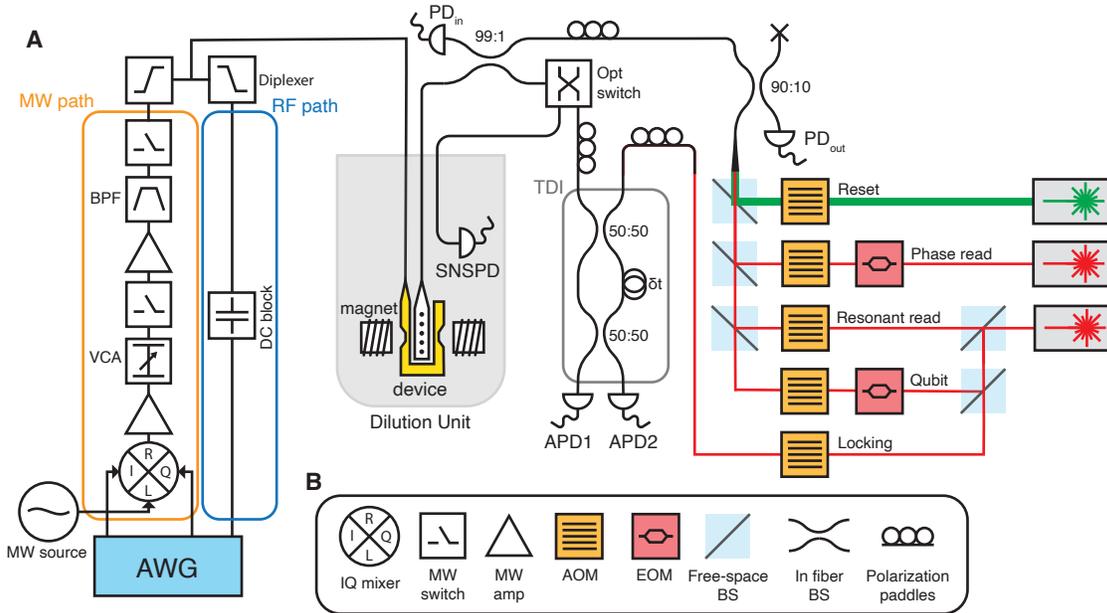


Fig. S1. (A) Diagram of the setup for optical and microwave control of our system. (B) Legend of components in (A).

The MW and RF chain is driven using a Zurich Instruments HDAWG8 2.4 GSa/s AWG that is also used to time the experimental sequence, switch the AOMs, and generate the qubit gate waveforms on the EOM. On the microwave path (from bottom to top in Fig. S1) are an IQ mixer (Marki MMIQ-0218L), a pre-amplifier (ZX60-06183LN+), a voltage controlled attenuator (RFVA T0218A50), a switch (CMCS0947A-C2), an amplifier (ZVE-3W-183+), a bandpass filter (ZVBP-11R7G-S+, 10.70 – 12.75 GHz), a high power switch (Qorvo QPC2040), and a diplexer (ZDSS-3G4G-S+). The microwave source used as the LO of the IQ mixer is an Agilent 83732B Synthesized Signal Generator and is set to 11.85 GHz at 14 dBm. The RF pulses are directly synthesized from the HDAWG and sent to the diplexer via a DC block (MCL15542 BLK-18-S+). The microwave and RF signals are sent into the dilution refrigerator via a coaxial

cable that is wirebonded to the coplanar waveguides on the diamond chip. All experimental sequences are run using `pylabnet`, a custom Python data-taking software framework (36), together with fast logic on the AWG.

3 Cavity-QED Model

We use the following expressions for the reflection, transmission, and scattering amplitudes of light for a two-level system coupled to a cavity (37):

$$r(\omega) = 1 - \frac{\kappa_{\text{in}}}{i(\omega - \omega_c) + \kappa_{\text{tot}}/2 + g^2/(i(\omega - \omega_a) + \gamma/2)}$$

$$t(\omega) = \frac{\sqrt{\kappa_{\text{in}}\kappa_{\text{out}}}}{i(\omega - \omega_c) + \kappa_{\text{tot}}/2 + g^2/(i(\omega - \omega_a) + \gamma/2)}$$

$$s(\omega) = \frac{\sqrt{\kappa_{\text{in}}\gamma}g}{(i(\omega - \omega_c) + \kappa_{\text{tot}}/2)(i(\omega - \omega_a) + \gamma/2) + g^2}$$

where $r(\omega)$, $t(\omega)$, and $s(\omega)$ are the reflection, transmission, and scattering amplitudes respectively. The transmitted component is defined to include both light that is coupled through the transmission port as well as light that is scattered off the cavity due to fabrication imperfections; both modes of losses are incorporated into κ_{out} . κ_{in} is the coupling rate of the cavity into the in-coupling port while $\kappa_{\text{tot}} = \kappa_{\text{in}} + \kappa_{\text{out}}$ is the total cavity decay rate. The scattered component is defined as the component that decays into free space due to spontaneous emission, and is characterized by the bare SiV emitter linewidth γ . Finally, $\omega_{c(a)}$ is the resonance frequency of the cavity (SiV) and g is the coupling rate of the SiV to the cavity. The cooperativity of the system is $C = \frac{4g^2}{\kappa_{\text{tot}}\gamma} = 1.6 \pm 0.1$. Fig. 1D in the main text shows the intensity and phase of the reflected light, where the phase is given by $\arg(r(\omega))$ and the intensity by $R(\omega) = |r(\omega)|^2$. The datapoints were fitted to obtain the parameters ω_a for the 4 transitions together with κ_{in} and g . The cavity frequency ω_c and linewidth κ_{tot} were separately fitted from the bare cavity reflection spectrum.

Parameter	Description	Fitted Value
ω_c	Cavity resonance frequency	406.610(1) THz
$\omega_{a,\uparrow_e\uparrow_n}$	Optical resonance frequency for $ \uparrow_e\uparrow_n\rangle$	406.678504(2) THz
$\omega_{a,\uparrow_e\downarrow_n}$	Optical resonance frequency for $ \uparrow_e\downarrow_n\rangle$	406.678537(2) THz
$\omega_{a,\downarrow_e\uparrow_n}$	Optical resonance frequency for $ \downarrow_e\uparrow_n\rangle$	406.679053(2) THz
$\omega_{a,\downarrow_e\downarrow_n}$	Optical resonance frequency for $ \downarrow_e\downarrow_n\rangle$	406.679022(2) THz
$\kappa_{\text{in}}/2\pi$	Coupling rate into cavity	202(12) GHz
$\kappa_{\text{tot}}/2\pi$	Total cavity linewidth	250(14) GHz
$g/2\pi$	SiV-cavity coupling rate	3.19(5) GHz
$\gamma/2\pi$	Bare SiV linewidth	100 MHz

Table S1: Numerical values of cavity and emitter parameters. The number in parentheses indicate the uncertainty in the last digits. All linewidths and rates are defined as the full-width at half-maximum values.

4 MW and RF gates

All relevant MW and RF transition frequencies as well as pulse parameters are listed in Table S2, and Rabi oscillations for the different transitions are shown in Fig. S2. The relatively large nuclear Rabi frequency is due to the strong electron-nuclear hyperfine interaction strength. The Rabi frequency Ω_e and π -pulse time τ_e for electron-flipping pulses need to satisfy $\Omega_e\tau_e = \pi$. However, in order to achieve a $C_n\text{NOT}_e$, we need to ensure that the off-resonant electron transition (i.e. when the nucleus is not in the state being conditioned on) is not flipped by the off-resonant driving. The off-resonant transition is detuned by the ^{29}Si -electron hyperfine coupling strength $A_{\parallel} = 2\pi \times 66.25$ MHz, so we require $\sqrt{A_{\parallel}^2 + \Omega_e^2}\tau_e = 2m\pi$ for any positive integer m in order for the electron to rotate back to its initial state. This constraint results in discrete choices of the Rabi frequency, namely $\Omega_e = A_{\parallel}/\sqrt{4m^2 - 1}$. We choose m to be as small as possible given the constraints on our maximal drive power ($m = 2$) to minimize the effect of electron qubit frequency diffusion (~ 1 MHz). For nuclear-flipping pulses, heating limitations force us to work at low Rabi frequencies $\Omega_n \ll \omega_{RF2} - \omega_{RF1}$, so off-resonant transition driving is negligible.

While the above choice of Rabi frequency ensures that the electron is not flipped by off-resonant driving, it still picks up a geometric phase from its rotation on the Bloch sphere, which can be detected when the nucleus is placed in a superposition and $C_n\text{NOT}_e$ is applied twice:

$$|\downarrow_e\rangle (|\downarrow_n\rangle + |\uparrow_n\rangle) \xrightarrow{2 \times C_n\text{NOT}_e} |\downarrow_e\rangle \left(e^{i\gamma_{res}} |\downarrow_n\rangle + e^{4i\gamma_{det}} |\uparrow_n\rangle \right)$$

where γ_{res} and γ_{det} refer to the geometric phase picked up by the electron during one closed loop around the resonant and detuned Rabi driving axes respectively, and the factors of 1 and 4 account for the fact that the off-resonant driving causes 4π rotation for every π rotation with

Parameter	Value
ω_{MW1}	12.00746 GHz
ω_{MW2}	12.07384 GHz
ω_{RF1}	29.636 MHz
ω_{RF2}	36.614 MHz
$\Omega_e/2\pi$	16.7 MHz
τ_e	30.0 ns
$\Omega_n/2\pi$	19.6 kHz (RF1), 24.2 kHz (RF2)
τ_n	25.5 μ s (RF1), 20.7 μ s (RF2)
$T_{1,e}$	2.9 s (0.1 K), 17 ms (4.3 K)
$T_{2,e}^*$	~ 5 μ s (0.1 K)
$T_{2,n}^*$	~ 5 ms (0.1 K)
$T_{2,e}$ with Hahn echo	78 μ s (0.1 K), 400 ns (4.3 K)
$T_{2,n}$ with Hahn echo	79 ms (0.1 K), 4.5 ms (4.3 K)
Magnetic field	0.394 T

Table S2: Microwave and RF driving parameters.

resonant driving. The geometric phase can be written as $\gamma = -\frac{S}{2}$ where S is the solid angle on the Bloch sphere subtended by the closed path. In our case, resonant and detuned driving cause the Bloch vector to be driven around the axes $(\Omega_e, 0, 0)$ and $(\Omega_e, 0, A_{\parallel})$ respectively, which gives the respective geometric phases of $\gamma_{res} = -\pi$ and $\gamma_{det} = -\pi \left(1 - \frac{\sqrt{15}}{4}\right)$. The relative phase Γ induced on the nuclear superposition as a result of applying $C_n\text{NOT}_e$ twice as described above would therefore be $\Gamma \equiv 4\gamma_{det} - \gamma_{res} \approx 0.87\pi$.

We verify this expression by performing a modified nuclear Ramsey sequence where we drive $C_n\text{NOT}_e$ pulses on the electron spin during the free evolution time. From Fig. S3, we measure that $\Gamma = 0.85\pi$ and 0.87π for $C_n\text{NOT}_e$ and $\overline{C_n\text{NOT}_e}$ respectively, which matches our predictions. More importantly, applying both CNOTs twice during the sequence leads to no phase shift, confirming that the geometric phases imparted cancel out. This allows us to neglect this phase as all relevant pulse sequences in our experiments have $C_n\text{NOT}_e$ and $\overline{C_n\text{NOT}_e}$ pulses paired appropriately such that this phase is canceled.

For the decoupled $C_e\text{NOT}_n$, the length of each individual RF pulse was chosen to be precisely an even number of periods. This guarantees that the average magnetic field due to the RF driving in any XY8 waiting window goes to zero, eliminating any extra phase induced on the electron from RF driving. While the pulse sequence for a decoupled $C_e\text{NOT}_n$ was shown in Fig. 2A of the main text, a decoupled $C_e\text{NOT}_n$ gate can be analogously implemented by switching the positions of the RF1 and RF2 pulses.

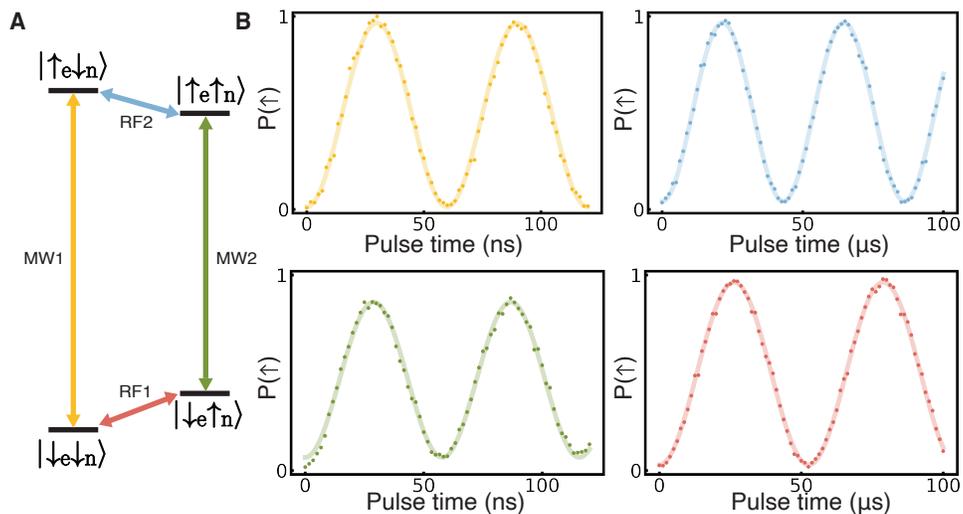


Fig. S2. (A) Qubit manifold energy level diagram of ^{29}SiV . Driving on each labeled transition induces a conditional rotation. (B) Rabi oscillations for each of the four transitions.

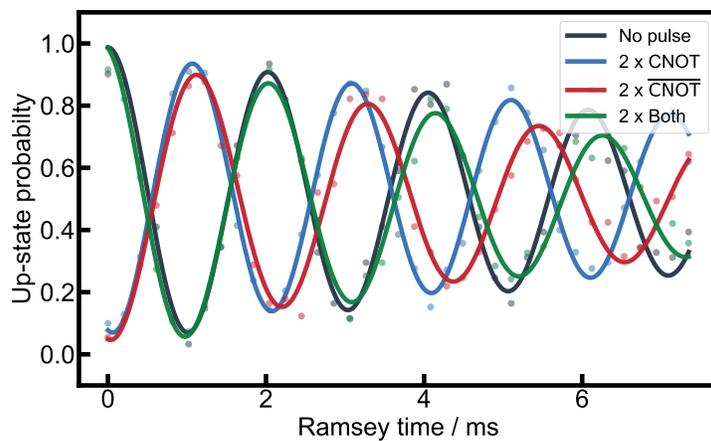


Fig. S3. Measurement of nuclear Ramsey oscillations with various pulses applied during the free precession time. Nuclear $\pi/2$ pulses were intentionally detuned by 500 Hz so that Ramsey fringes can be observed.

5 ^{13}C Characterization

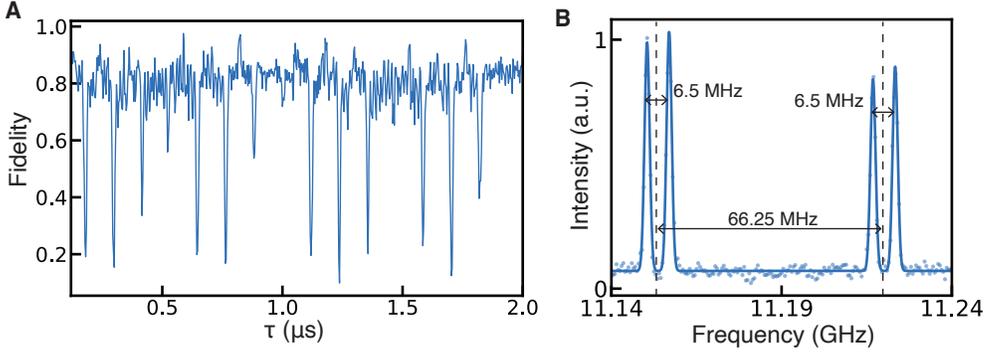


Fig. S4. (A) XY8 decoupling sequence on the electron, sweeping the inter-pulse time 2τ . (B) Optically detected magnetic resonance (ODMR) spectrum of the electron with all nuclear spins uninitialized. The larger splitting (66.25 MHz) is due to the ^{29}Si spin while the smaller splitting (6.5 MHz) is due to the ^{13}C spin.

In addition to the strongly-coupled ^{29}Si nucleus explored in this paper, our device also contains a comparatively more weakly coupled ^{13}C nucleus located nearby in the lattice. The coupling is however large enough to induce a splitting of measured electron transitions (Fig. S4B). The ^{13}C nuclear spin was not used in this paper; it is characterized here and was initialized in all of our experiments so that its effect can be neglected. The ^{13}C -SiV Hamiltonian is similar to the ^{29}Si -SiV Hamiltonian, though the magnetic field is not necessarily aligned with the hyperfine interaction, resulting in a parallel and perpendicular component of the coupling matrix \mathbf{A} (38):

$$\mathcal{H} = \gamma_e S_z + \gamma_n I_z + A_{\parallel} S_z I_z + A_{\perp} S_z I_x$$

To evaluate the electron coupling to the nearby ^{13}C , we performed a sweep of the inter-pulse time of an electron XY8 sequence (Fig. S4A). Fitting of this signal indicates a parallel coupling component of $A_{\parallel} = 2\pi \times 3.2$ MHz and a perpendicular coupling of $A_{\perp} = 2\pi \times 0.36$ MHz. The ^{13}C -flipping RF transitions, which we can drive directly, are 1.041 MHz and 7.475 MHz when the electron spin is in the down and up state respectively. We achieve a Rabi frequency of the transition of 5.23 kHz, corresponding to a π time of 95.6 μs . We use this to initialize the ^{13}C at the beginning of every experimental sequence by swapping the electron state onto the ^{13}C (with $C_{13\text{C}}\text{NOT}_e - C_e\text{NOT}_{13\text{C}} - C_{13\text{C}}\text{NOT}_e$), and then confirm whether the initialization was successful by performing a Ramsey sequence on the electron to precisely determine its transition frequency. This allows for a readout of the ^{13}C state since the ^{13}C nucleus induces a spin-dependent shift on the electron transition frequency.

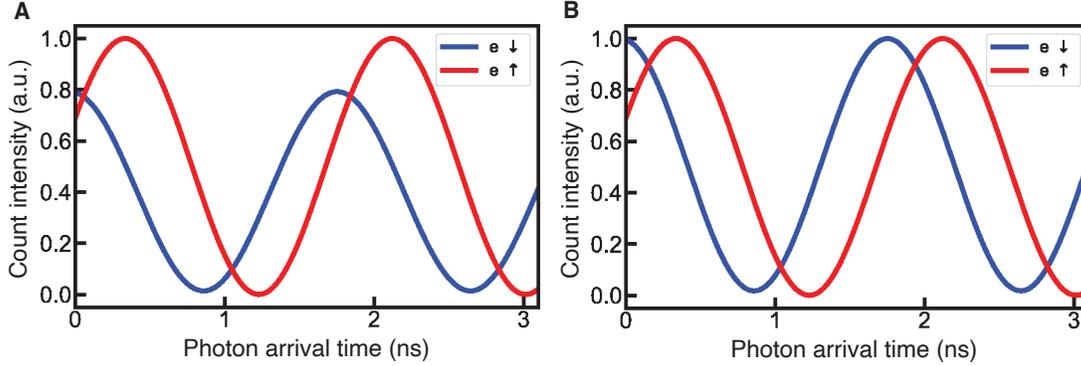


Fig. S5. (A) Simulated arrival time probability distributions for both electron states before and (B) after normalization. The normalized distribution as a function of time is called $P_{\downarrow/\uparrow}(t)$. The distributions are shifted in time as the reflected phase difference $\Delta\phi$ between the two measurement sideband frequencies depend on the electron state.

6 Phase Readout of Electron State

We implement the phase-based readout of the electron state by using an EOM to frequency-shift a carrier laser into two sidebands located 300 MHz above and below the carrier (ω_r and ω_b , magenta and cyan lines in Fig. 1D). Both sideband components travel along the same path, negating the need for any phase stabilization, and are reflected off the device before being detected on a SNSPD with a time resolution of 70 ps (Fig. 1E, F). We make use of the fact that there is a sizeable electron state dependence of the reflected phase difference between the two sideband laser frequencies. When these two sidebands are reflected off the cavity, they experience a phase of $\phi_{\downarrow/\uparrow}^b$ and $\phi_{\downarrow/\uparrow}^r$ respectively, with the phase acquired dependent on the SiV electron state. The value of these phases are given by $\arg(r(\omega_b))$ and $\arg(r(\omega_r))$, where $r(\omega)$ is the complex reflection amplitude at a laser frequency ω .

When these sidebands are sent onto a single photon detector, the photon arrival time probability distribution function is a beat signal at their frequency difference of 600 MHz with a phase of $\Delta\phi_{\downarrow/\uparrow} \equiv (\phi_{\downarrow/\uparrow}^b - \phi_{\downarrow/\uparrow}^r)$ depending on the electron spin state. The phase difference for the two electron states ($\Delta\phi_{\downarrow} - \Delta\phi_{\uparrow}$) is observed as a phase shift between the arrival time distributions as seen in Fig. S5A. The experimentally measured beats in Fig. 1F are fit using sinusoidal oscillations at 600 MHz with a phase shift computed from the cavity-QED model and the relative amplitudes calculated from the reflectances $|r(\omega_b)|^2$ and $|r(\omega_r)|^2$. A small background signal in the oscillations is modeled as leakage of the carrier frequency through the EOM.

The fact that the photon arrival time depends on the electron state means that detection of a photon at a given time provides us with information on the electron state, with the information content proportional to the likelihood ratio of the probability distribution functions. We will quantify this information gain using Bayesian analysis and explain how we use this as a readout protocol. We first note that there is an overall scaling difference between the two probability

distributions. This is because the spin-up state has a higher reflectance at the phase readout frequencies (Fig. 1D), leading to a larger number of counts by 14%. The curves are separately normalized (Fig. S5B) and we call these the arrival time probability functions $P_{\uparrow/\downarrow}(t)$.

During the phase readout procedure, the readout lasers are turned on for a fixed period of time and a variable number N of photons are collected that arrive at times $\{t_1, t_2, \dots, t_N\}$. Our initial estimate for the probability that the electron is in spin state \downarrow is chosen to be the uniform prior $p_{\downarrow}^{(0)} = 0.5$. We then repeatedly apply Bayes' rule to update our probability estimate $p_{\downarrow}^{(i)}$ after the arrival of photon i at time t_i . We note that in reality the analysis is done in postprocessing after all N photons have been collected, but it is equivalent to consider performing the analysis in an iterative manner for each individual arriving photon. The posterior probability estimate is:

$$p_{\downarrow}^{(i)} = \frac{p_{\downarrow}^{(i-1)} P_{\downarrow}(t_i)}{p_{\downarrow}^{(i-1)} P_{\downarrow}(t_i) + p_{\uparrow}^{(i-1)} P_{\uparrow}(t_i)} \quad (\text{S1})$$

where $p_{\uparrow}^{(i)} \equiv 1 - p_{\downarrow}^{(i)}$. After incorporating the information from the arrival times of all N photons, we apply Bayes' rule one more time, this time using the fact that the two electron states have different overall reflectivities and thus they have a different mean number of reflected photons $\bar{n}_{\uparrow/\downarrow}$ that we expect to detect during the readout window. Our final estimate for the probability that the electron is in the spin-down state is therefore:

$$p_{\downarrow} = \frac{p_{\downarrow}^{(N)} \text{Po}(\bar{n}_{\downarrow}, N)}{p_{\downarrow}^{(N)} \text{Po}(\bar{n}_{\downarrow}, N) + p_{\uparrow}^{(N)} \text{Po}(\bar{n}_{\uparrow}, N)} \quad (\text{S2})$$

Here, $\text{Po}(\bar{n}, N)$ is the probability mass function of a Poisson distribution with mean \bar{n} to have N photons detected. This last factor justifies our decision to separately normalize each electron state's probability distribution previously when defining $P_{\uparrow/\downarrow}(t)$, since the probability of detecting N photons at times $\{t_1, t_2, \dots, t_N\}$ can be decomposed into (I) the probability of detecting N photons in total (accounted for in Eqn. S2) multiplied by (II) the probability of detecting photons at times $\{t_1, t_2, \dots, t_N\}$ conditioned on the fact that we have detected N photons in total (accounted for in Eqn. S1 by using the normalized distributions).

At the end of this process, the electron is determined to be in the down state if $p_{\downarrow} > 0.5$ and in the up state otherwise. In order to reduce the readout error rate, we can perform postselection on the posterior probabilities by determining that the electron is in the down (up) state only if $p_{\downarrow} > 1 - \varepsilon$ ($p_{\downarrow} < \varepsilon$) for some stricter choice of threshold $\varepsilon < 0.5$. All readout attempts with $\varepsilon \leq p_{\downarrow} \leq 1 - \varepsilon$ would be discarded as failed measurements. Such a postselection protocol could be beneficial in scenarios where the accuracy of readout is of paramount importance and where the state preparation can be repeated if necessary (e.g. in a quantum repeater). In the following discussions and in the main text, we fix $\varepsilon = 0.2$ for simplicity.

Fig. S6 shows how the readout error decreases as we increase the mean number of photons \bar{n} collected during phase readout window by increasing the readout time. The errors are averaged

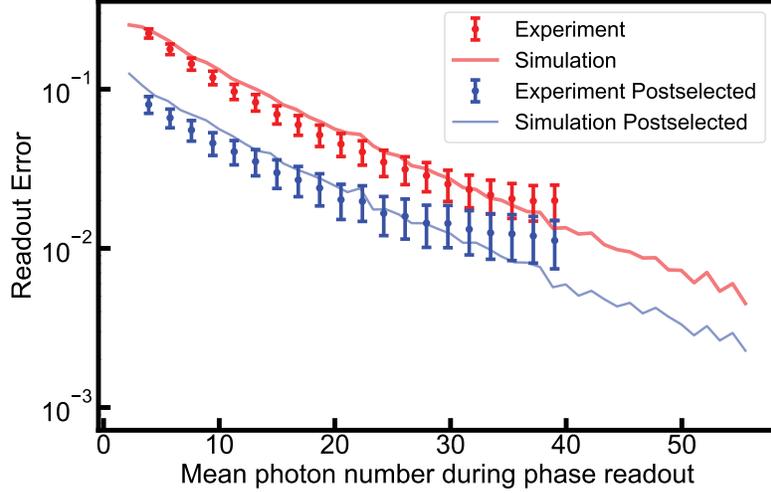


Fig. S6. Comparison of phase readout error rates against the mean number of photons collected during each readout window without (red) and with (blue) postselection at 80% confidence.

over both spin states. The readout error rate is estimated experimentally by initializing in an X -basis state and then read out with the phase readout followed by the resonant readout, with the latter assumed to be the true electron state. A simulation of the expected error scaling was performed by generating random photon arrival time samples from the experimentally-measured arrival time distribution.

This error rate scaling can be combined with our measurement of nuclear decoherence from electron readout (Fig. S8) to characterize the nuclear dephasing rate of the phase readout protocol compared to resonant readout. Fig. S7 shows that for every desired electron readout fidelity of our choice, the phase read procedure (both with and without postselection) leads to less nuclear state decoherence. For instance, if we desire an electron readout with 5% error, the electron can be read 14 times using the postselected phase read before we suffer a $1/e$ decoherence of the nuclear state. In contrast, the resonant readout can only be performed 1.8 times before the same amount of nuclear decoherence sets in, which is a factor of 8 lower.

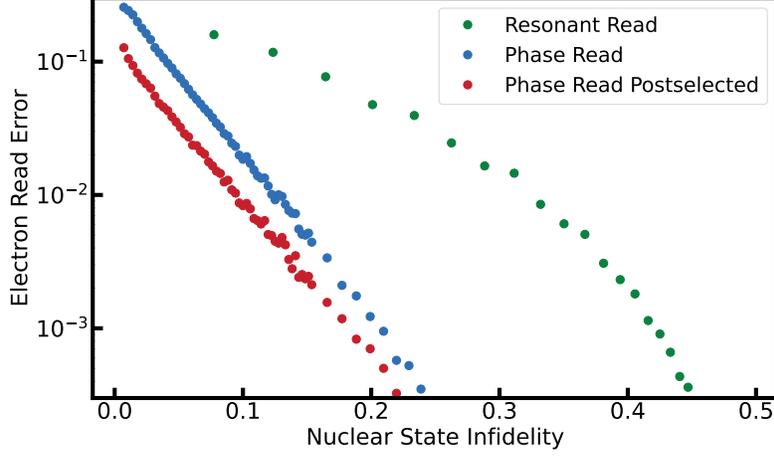


Fig. S7. Scaling of simulated electron readout error against nuclear state infidelity caused by the readout. Nuclear state infidelity is defined in a similar manner to Fig. S8, i.e. it is 0 when we have perfect coherence of a X nuclear superposition state and 0.5 when the nuclear phase is completely decohered by the electron readout.

7 Laser-Induced Nuclear Decoherence

Due to the difference in the hyperfine coupling between the ground and excited states, the optical transition frequency is shifted depending on the state of the ^{29}Si nuclear spin, causing the reflection amplitude and phase spectra to be different for different nuclear states (Fig. 1D in main text). As a result, the scattered readout photons acquire information about the nuclear state, leading to decoherence of the nuclear superposition once the photon states are traced over. To model the decoherence process, consider a coherent state $|\alpha\rangle$ incident on a uniform superposition of the nuclear state $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ during a readout step. Here, we assume the electron is in an initialized state and we will omit its state in the following expressions. After scattering the coherent state off the SiV, the light that is reflected, transmitted, and scattered off the cavity can be expressed as:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} (|\downarrow\rangle \otimes |R_{\downarrow}\alpha\rangle \otimes |T_{\downarrow}\alpha\rangle \otimes |S_{\downarrow}\alpha\rangle + |\uparrow\rangle \otimes |R_{\uparrow}\alpha\rangle \otimes |T_{\uparrow}\alpha\rangle \otimes |S_{\uparrow}\alpha\rangle)$$

where $R_{\uparrow/\downarrow}, T_{\uparrow/\downarrow}, S_{\uparrow/\downarrow}$ refer to the reflection, transmission, and scattering coefficients for the cavity with the ^{29}Si nuclear spin in the state indicated in the subscript. We can then describe the

system by the density matrix:

$$\begin{aligned} \rho_1 = |\psi_1\rangle\langle\psi_1| = \frac{1}{2} (& |\downarrow\rangle\langle\downarrow| \otimes |R_\downarrow\alpha\rangle\langle R_\downarrow\alpha| \otimes |T_\downarrow\alpha\rangle\langle T_\downarrow\alpha| \otimes |S_\downarrow\alpha\rangle\langle S_\downarrow\alpha| + \\ & |\uparrow\rangle\langle\uparrow| \otimes |R_\uparrow\alpha\rangle\langle R_\uparrow\alpha| \otimes |T_\uparrow\alpha\rangle\langle T_\uparrow\alpha| \otimes |S_\uparrow\alpha\rangle\langle S_\uparrow\alpha| + \\ & |\downarrow\rangle\langle\uparrow| \otimes |R_\downarrow\alpha\rangle\langle R_\uparrow\alpha| \otimes |T_\downarrow\alpha\rangle\langle T_\uparrow\alpha| \otimes |S_\downarrow\alpha\rangle\langle S_\uparrow\alpha| + \\ & |\uparrow\rangle\langle\downarrow| \otimes |R_\uparrow\alpha\rangle\langle R_\downarrow\alpha| \otimes |T_\uparrow\alpha\rangle\langle T_\downarrow\alpha| \otimes |S_\uparrow\alpha\rangle\langle S_\downarrow\alpha|) \end{aligned}$$

The transmitted and scattered light is lost to our experimental setup, while the reflected light is collected and measured by our photon counters (up to collection loss). However, as we do not condition any of the subsequent experimental procedures on the measurement result, we can trace over the reflected photon states as well. Tracing over all the photon states gives the resulting nuclear state:

$$\begin{aligned} \rho_1 &= \frac{1}{2} \left[\begin{array}{cc} \langle R_\downarrow\alpha|R_\downarrow\alpha\rangle \langle T_\downarrow\alpha|T_\downarrow\alpha\rangle \langle S_\downarrow\alpha|S_\downarrow\alpha\rangle & \langle R_\downarrow\alpha|R_\uparrow\alpha\rangle \langle T_\downarrow\alpha|T_\uparrow\alpha\rangle \langle S_\downarrow\alpha|S_\uparrow\alpha\rangle \\ \langle R_\uparrow\alpha|R_\downarrow\alpha\rangle \langle T_\uparrow\alpha|T_\downarrow\alpha\rangle \langle S_\uparrow\alpha|S_\downarrow\alpha\rangle & \langle R_\uparrow\alpha|R_\uparrow\alpha\rangle \langle T_\uparrow\alpha|T_\uparrow\alpha\rangle \langle S_\uparrow\alpha|S_\uparrow\alpha\rangle \end{array} \right] \\ &\equiv \frac{1}{2} \begin{bmatrix} 1 & \rho_{1,\downarrow\uparrow} \\ \rho_{1,\downarrow\uparrow}^* & 1 \end{bmatrix} \end{aligned}$$

where the matrix is written in the nuclear state basis $\{|\downarrow\rangle, |\uparrow\rangle\}$. The off-diagonal term $\rho_{1,\downarrow\uparrow}$ measures the coherence of the state and is given by:

$$\rho_{1,\downarrow\uparrow} = \prod_{C \in \{R,T,S\}} \langle C_\downarrow\alpha|C_\uparrow\alpha\rangle = \prod_{C \in \{R,T,S\}} \exp \left[-\frac{1}{2} (|C_\downarrow\alpha|^2 + |C_\uparrow\alpha|^2) + (C_\downarrow\alpha)^*(C_\uparrow\alpha) \right]$$

where C runs over all three paths of the light and the final term arises from the inner product properties of coherent states. The term in the product can be reformulated as:

$$\exp \left[-\frac{1}{2} (|C_\downarrow\alpha - C_\uparrow\alpha|^2) + \frac{1}{2} ((C_\downarrow\alpha)^*(C_\uparrow\alpha) - (C_\downarrow\alpha)(C_\uparrow\alpha)^*) \right]$$

We can add further clarity by noting that $a^*b - ab^*$ is purely imaginary for all complex numbers a and b since it can be rewritten as $2i(\text{Re}(a)\text{Im}(b) - \text{Im}(a)\text{Re}(b))$, and thus we can rewrite $(C_\downarrow\alpha)^*(C_\uparrow\alpha) - (C_\downarrow\alpha)(C_\uparrow\alpha)^*$ as $2i\phi$ for some real ϕ that depends on α . This allows us to simplify:

$$\rho_{1,\downarrow\uparrow} = \prod_{C \in \{R,T,S\}} \exp \left[-\frac{1}{2} (|C_\downarrow\alpha - C_\uparrow\alpha|^2) \right] e^{i\phi}$$

separating the effect of the light into a pure loss of coherence and a pure rotation similar to a Stark shift.

The coherence can then be measured by measuring the expectation value of the nuclear state

in the X -basis, which gives:

$$\langle X_1 \rangle = \text{Tr } X \rho_1 = \frac{\rho_{1,\downarrow\uparrow} + \rho_{1,\downarrow\uparrow}^*}{2} = \text{Re } \rho_{1,\downarrow\uparrow}$$

This is the experimental quantity that we measure to determine the decoherence induced on the nucleus by scattering photons.

We can cancel out the coherent phase rotation on the nucleus by performing a π phase flip on the nucleus between two separate coherent state packets $|\alpha\rangle$, in a manner similar to that in a spin echo experiment. This effectively performs the state readout in two separate windows. The resulting state can be derived using an almost identical process, giving:

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle \otimes |R_\downarrow\alpha\rangle \otimes |T_\downarrow\alpha\rangle \otimes |S_\downarrow\alpha\rangle \otimes |R_\uparrow\alpha\rangle \otimes |T_\uparrow\alpha\rangle \otimes |S_\uparrow\alpha\rangle + \\ |\downarrow\rangle \otimes |R_\uparrow\alpha\rangle \otimes |T_\uparrow\alpha\rangle \otimes |S_\uparrow\alpha\rangle \otimes |R_\downarrow\alpha\rangle \otimes |T_\downarrow\alpha\rangle \otimes |S_\downarrow\alpha\rangle)$$

where the first (second) three photonic kets refer to photons scattered in the first (second) time window. Taking the density matrix as before:

$$\rho_2 = \frac{1}{2} \begin{bmatrix} 1 & \rho_{2,\downarrow\uparrow} \\ \rho_{2,\downarrow\uparrow}^* & 1 \end{bmatrix}$$

with

$$\rho_{2,\downarrow\uparrow} = \prod_{C \in \{R,T,S\}} |\langle C_\downarrow\alpha | C_\uparrow\alpha \rangle|^2 = \prod_{C \in \{R,T,S\}} \exp \left[-|C_\downarrow\alpha - C_\uparrow\alpha|^2 \right]$$

which gives rise to a pure exponential decay with the oscillations canceled out. As before, the experimental observable is the X expectation value which is given by:

$$\langle X_2 \rangle = \text{Tr } X \rho_2 = \text{Re } \rho_{2,\downarrow\uparrow} = \rho_{2,\downarrow\uparrow}$$

All desired dependencies of the nuclear decoherence rate on the readout parameters can be derived from these expressions for the X expectation value. In particular, the laser readout frequency affects the reflection, transmission and scattering coefficients R, T, S according to the Jaynes-Cummings model, while the readout length and laser intensity affects the coherence state α via the mean photon number N in each readout packet with $\langle N \rangle = |\alpha|^2$.

We verify this model of decoherence with the experimental sequence shown in Fig. S8. The ^{29}Si nuclear spin is first prepared in a superposition with a $\frac{\pi}{2}$ pulse, and the readout laser is applied in either one or two readout windows separated by a π pulse on the nucleus. The nuclear spin is then read out in the X basis by applying a $\frac{\pi}{2}$ pulse followed by reading out via the electron. Fig. S8A shows that the nuclear coherence exhibits oscillations from the coherent accumulation of phase and an exponential decay from the decoherence process predicted above, while applying the laser pulse in a symmetrical manner around the π pulse cancels out the

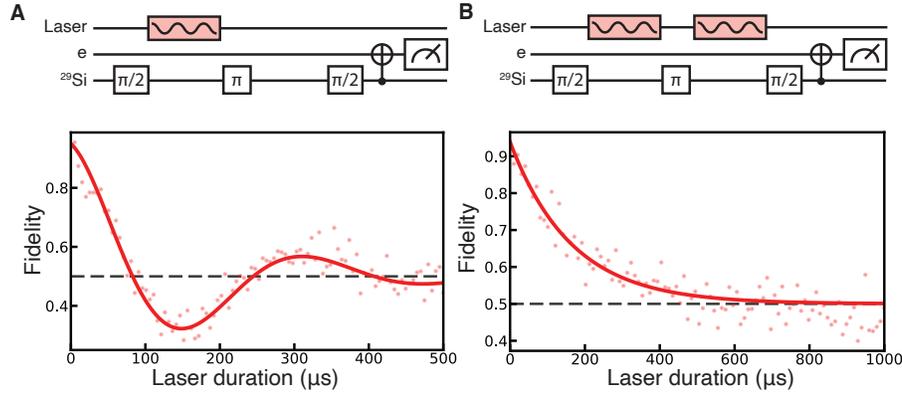


Fig. S8. (A) Comparison of nuclear decoherence rate against laser duration when a laser at the resonant readout frequency is applied in an asymmetric and (B) symmetric manner during a spin echo measurement. The resulting coherence of the nucleus is read out in the X basis by applying a $\pi/2$ pulse.

oscillation and highlights only the laser-induced decoherence (Fig. S8B). The plots are fit with the expressions derived for $\langle X_1 \rangle$ and $\langle X_2 \rangle$ respectively using the reflection, transmission, and scattering coefficients computed from the cavity-QED model.

This model can be further explored by sweeping the frequency of the laser being applied while keeping the laser length constant (Fig. S9). The amount of decoherence experienced by the nucleus at a given laser frequency depends on the electron state. The black dashed line indicates the resonant readout frequency, i.e. where the SiV contrast is largest, and is also the frequency used above in Fig. S8. The fact that it lies so close to a minimum of the nuclear coherence confirms the argument in the main text that resonant readout, while being a high-fidelity electron state readout method, also reads out the nuclear state and leads to rapid decoherence of any nuclear superposition. In contrast, the frequencies used for the phase readout (cyan and magenta lines) lead to much weaker nuclear decoherence as they do not reveal much information about the nuclear state in both amplitude and phase (Fig. 1D), making them suitable for repeatedly reading and resetting the electron while using the nucleus as a long-lived memory.

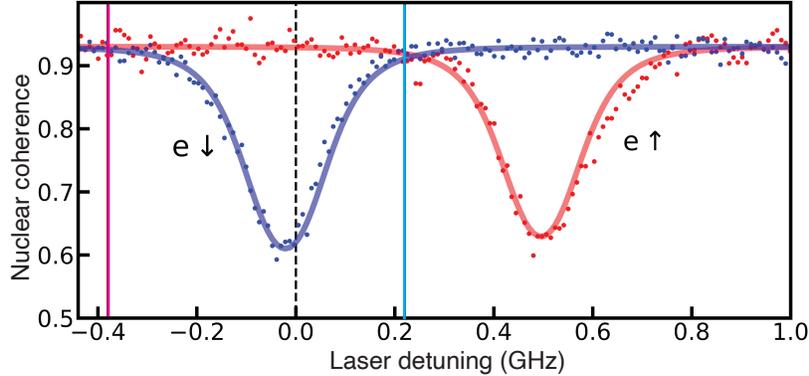


Fig. S9. Nuclear decoherence from symmetric application of a laser during spin echo process as laser frequency is varied. The amount of nuclear decoherence experienced depends on the electron state. The dashed vertical line is the laser position for resonant readout; the cyan and magenta vertical lines are the frequencies for phase readout.

8 Gate fidelity decomposition

Here, we characterize the fidelities of the various gates and constituent operations in our system. Initialization of the electron spin in the spin-down (spin-up) state is performed by applying a resonant readout laser pulse and checking if the reflected photon counts are below (above) a set initialization threshold. If not, a π pulse is applied to the electron and the readout is repeated until the desired state is measured. The initialization threshold for initializing in the spin-down (spin-up) state is lower (higher) than the readout threshold used for state determination, which allows for higher initialization fidelity and mitigates spectral diffusion by only allowing the initialization to succeed when there is sufficient reflection contrast between the spin states. By initializing the electron the down state and subsequently reading it out, we determine that our combined initialization and readout fidelity in the down state is $99.5 \pm 0.1\%$ (Fig. S10A). Readout of the nuclear state is performed with a CNOT followed by resonant electron readout (Fig. S10C).

To initialize both the ^{29}Si and ^{13}C nuclear spins, we initialize the electron in the spin-down state and swap the state onto the ^{29}Si before repeating the same steps for the ^{13}C . To confirm the success of the swap gates, we perform a Ramsey sequence on the lowest-frequency electron transition (corresponding to the $^{29}\text{Si}\downarrow\ ^{13}\text{C}\downarrow$ state) to determine if the entire system is in the correct state ($e\downarrow\ ^{29}\text{Si}\downarrow\ ^{13}\text{C}\downarrow$), where successful initialization is indicated by the lack of rapid Ramsey fringes that would be observed if the system were in a different state (which would have a detuned transition frequency). We then do an electron ODMR sequence to determine the nuclear populations and find that the ^{29}Si is correctly initialized $99.5 \pm 0.1\%$ of the time and the ^{13}C is correctly initialized $93.8 \pm 1.5\%$ of the time (Fig. S10B).

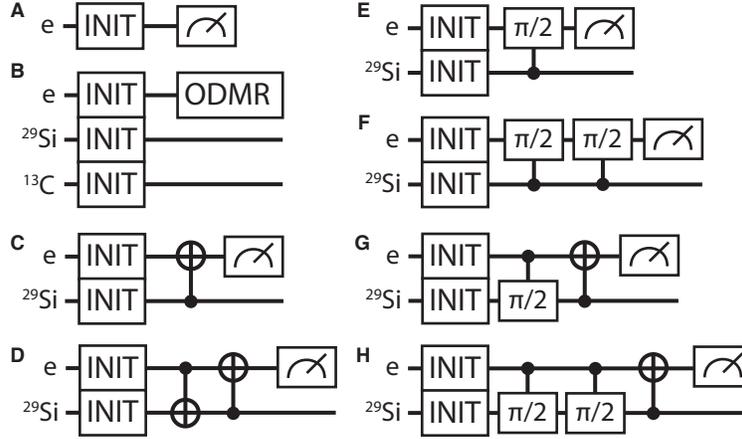


Fig. S10. (A) through (H): sequences to evaluate the fidelities of individual sequence components as described in the text. All initialization steps set the spins into the spin-down states, and all CNOT gates are conditioned on the control spin being in the spin-down state.

Sequence	P(\uparrow)	Deduced Gate Fidelity
C	$99.4 \pm 0.1\%$	$C_n\text{NOT}_e$ fidelity: 0.999 ± 0.001
D	$2.5 \pm 0.2\%$	$C_e\text{NOT}_n$ fidelity: 0.980 ± 0.002
E	$50.6 \pm 0.6\%$	$\sqrt{C_n\text{NOT}_e}$ fidelity: 0.988 ± 0.002
F	$97.1 \pm 0.2\%$	
G	$49.7 \pm 0.6\%$	$\sqrt{C_e\text{NOT}_n}$ fidelity: 0.990 ± 0.002
H	$98.0 \pm 0.2\%$	

Table S3: Outcomes of sequences shown in Fig. S10 with corresponding deduced gate fidelities.

We can now characterize the fidelities of the various qubit operations using the circuits as shown in Figs. S10C-H and we summarize the results in Table S3. These gate fidelities can be combined with the overall fidelity of the swap-hold-read sequence as performed in Fig. 2D to find that the decoupled $C_e\text{NOT}_n$ has a fidelity of $93.7 \pm 0.7\%$ (Table S4). The fidelity loss is mainly due to electron decoherence during the decoupling, as well as cumulated $C_n\text{NOT}_e$ and $C_n\text{NOT}_e$ gate infidelities.

Step	Fidelity
Initialization	0.995 ± 0.001
$\sqrt{C_n \text{NOT}_e}$	0.988 ± 0.002
Decoupled $C_e \text{NOT}_n$	0.937 ± 0.007
$C_n \text{NOT}_e$	0.999 ± 0.001
$C_e \text{NOT}_n$	0.980 ± 0.002
$\sqrt{C_e \text{NOT}_n}$	0.990 ± 0.002
$C_n \text{NOT}_e$	0.999 ± 0.001
Readout	0.995 ± 0.001
Total measured fidelity	0.891 ± 0.006
Total estimated fidelity	0.891 ± 0.008

Table S4: Fidelity decomposition for the Swap-Hold-Read sequence. The predicted fidelity is from the product of the fidelities of all steps.

9 Spin-photon gates operation and measurement

The initial photonic state for all spin-photon entangling gates is prepared as weak coherent pulses using an EOM. To limit the occurrence of two-photon events that would lower the fidelity of the final spin-photon entangled state, we work with pulses with average photon number 5×10^{-3} . The early and late photon timebins are separated by $\delta t = 143.5$ ns and have a Gaussian pulse shape of width 20.8 ns. The pulse width is chosen to be as wide as possible while still fully fitting within the separation time together with the microwave gates. This minimizes the spectral width of the photons, resulting in a better optical contrast. For the 4.3K operation of the PHONE gate, the photon pulses are shortened to square pulses of width 16.7 ns so that the $C_n \text{NOT}_e$ and $C_n \text{NOT}_e$ pulse pairs can be placed closer together. This is done to limit decoherence from the electron by lowering the time it is in a superposition state.

Measurement of the final entangled state is done by individually measuring the spin and the photon in different bases. To measure the photon in the Z basis, the photon arrival time is recorded to determine if it is in the early or late timebin. To measure the photon in the X and Y bases, we use an in-fiber TDI with a delay line equal to the photon spacing 143.5 ns corresponding to a fiber length of 43.05 m. The exact length of the delay line is adjusted with a piezoelectric hollow cylinder around which the fiber is wrapped. The TDI is locked every ~ 5 s by sending CW locking light through the locking port (Fig. S1A) and adjusting the delay line length to equalize counts on both outputs of the TDI. To measure in a specific basis, the frequency difference between the locking light and the photonic qubits is adjusted by changing the MW frequency sent to the locking light AOM (effectively changing the lock point) so that maximum photonic qubit contrast is achieved at the output, i.e. $|+X\rangle$ (or $|+Y\rangle$) results in clicks at APD1 only, and $| -X\rangle$ (or $| -Y\rangle$) results in clicks at APD2 only.

The measurements are combined to calculate the overlap fidelity with the $|\Phi^+\rangle = (|e \downarrow\rangle +$

$|l \uparrow\rangle\rangle/\sqrt{2} \equiv (|00\rangle + |11\rangle)/\sqrt{2}$ Bell state. For simplicity, here we denote the $|e\rangle$ and $|\downarrow\rangle$ states by $|0\rangle$ and the $|l\rangle$ and $|\uparrow\rangle$ states by $|1\rangle$, where the spin states can correspond to either to the electron or nuclear spin depending on the entangled state that we are preparing. The overlap fidelity with a given two-qubit density matrix ρ is given by:

$$F = \langle \Phi^+ | \rho | \Phi^+ \rangle = \frac{1}{2}(\rho_{11} + \rho_{44} + 2 \operatorname{Re}(\rho_{14}))$$

We want to express the above density matrix entries in terms of experimental observables. When the two-qubit system is measured in the XX , YY , and ZZ bases, the probability of obtaining the various measurement outcomes $(i, j) \in \{0, 1\}^2$ are:

$$\begin{aligned} p_{zz}^{ij} &= \langle ij | \rho | ij \rangle \\ p_{xx}^{ij} &= \langle ij | (\sqrt{X}_{\text{spin}} \otimes \sqrt{X}_{\text{photon}}) \rho (\sqrt{X}_{\text{spin}} \otimes \sqrt{X}_{\text{photon}}) | ij \rangle \\ p_{yy}^{ij} &= \langle ij | (\sqrt{Y}_{\text{spin}} \otimes \sqrt{Y}_{\text{photon}}) \rho (\sqrt{Y}_{\text{spin}} \otimes \sqrt{Y}_{\text{photon}}) | ij \rangle \end{aligned}$$

where the operators X and Y describe a π -rotation around the x and y -axes respectively: $\sqrt{X} = \exp\{i\pi S_x/2\hbar\}$, $\sqrt{Y} = \exp\{i\pi S_y/2\hbar\}$ with Pauli spin-1/2 matrices S_x and S_y . The measurement outcome probabilities can then be combined into the following linear combinations that relate to the relevant entries in the density matrix:

$$\begin{aligned} P_{zz} &\equiv p_{zz}^{00} + p_{zz}^{11} = \rho_{11} + \rho_{44} \\ P_{xx} &\equiv p_{xx}^{00} + p_{xx}^{11} - p_{xx}^{01} - p_{xx}^{10} = 4 \operatorname{Re}(\rho_{14}) + 4 \operatorname{Re}(\rho_{23}) \\ P_{yy} &\equiv p_{yy}^{01} + p_{yy}^{10} - p_{yy}^{00} - p_{yy}^{11} = 4 \operatorname{Re}(\rho_{14}) - 4 \operatorname{Re}(\rho_{23}) \end{aligned}$$

so that we are finally left with an explicit definition of the fidelity in terms of our measurement outcomes:

$$F = \frac{1}{2}P_{zz} + \frac{1}{4}P_{xx} + \frac{1}{4}P_{yy}$$

A decomposition of the sources of errors for spin-photon gates is given in Table S5. Note that the difference in SiV contrast between the electron-photon and PHONE gates is due to the nuclear state-dependence of the optical transition frequency (Fig. 1D) and thus contrast (Fig. S11). This means that the maximum achievable optical contrast for the PHONE gate will be lowered when we average over both nuclear spin states.

To determine the spin-photon heralding efficiency, we perform a spin echo sequence on the SiV electron. During the echo wait time, we apply a laser pulse of variable duration. We collect the reflected photons from this laser pulse and measure the state of the electron at the end of the echo sequence. Similar to the effect on the nucleus, reading out the electron during a spin echo dephases the superposition and causes the final electron state to exponentially decay towards a completely mixed state as a function of the average photon number arriving at the SiV. This gives us an estimate for N , the average number of photons that have interacted with the SiV, as

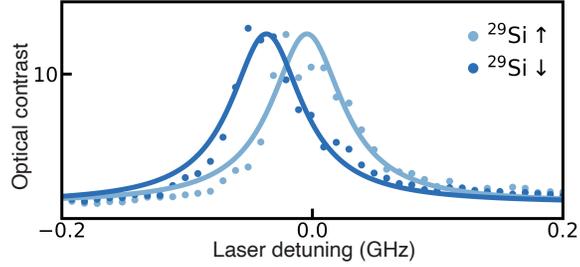


Fig. S11. Frequency-dependent contrast between the electron spin states for the two nuclear spin states.

a function of laser pulse time.

On the other hand, the number of reflected photons collected N_c during the sequence increases linearly with N . By plotting N_c against N for various laser pulse times, we obtain an estimate for the slope η , which represents the efficiency with which photons that have interacted with the SiV are able to be collected in the photodetectors, which is also the spin-photon heralding efficiency. Since this measurement was performed with SNSPDs with quantum efficiency near ~ 1 (15), we neglect this as a source of loss. We find that the efficiency $\eta = 0.143$. A major source of loss is the low reflectivity of the SiV (~ 0.275 , averaged over both electron states), while the remainder can be attributed to the coupling efficiency of the tapered fiber from the nanophotonic device ($\sim 0.6 - 0.7$), as well as losses in splicing points and in the 99:1 in-fiber beamsplitter along the path.

Measurement Basis	Electron-photon		PHONE	
	ZZ	XX, YY	ZZ	XX, YY
Temperature	0.1 K (1.5 K)		0.1 K (4.3 K)	
Initialization	0.995		0.995 (0.970)	
MW gates	0.99		0.94 (0.89)	
SiV contrast	0.96		0.94	
Dark counts	0.997		0.997	
2-photon events	0.998		0.998	
Readout	0.995		0.995	
TDI contrast	N/A	0.967	N/A	0.967
TDI lock and polarization drift	N/A	0.98	N/A	0.98
T_2 dephasing	N/A	0.997	N/A	0.997 (0.680)
Total estimated fidelity	0.91 (0.91)		0.85 (0.66)	
Total measured fidelity	0.91 ± 0.02 (0.90 ± 0.01)		0.85 ± 0.02 (0.66 ± 0.02)	

Table S5: Fidelity decomposition for all spin-photon gates. Parameters that take on a different value at temperatures of 1.5 K and 4.3 K for the electron-photon and PHONE gates respectively are shown in parentheses.

10 Temperature Robustness with SiV Strain

Within our experimental temperature regime of 100 mK to 4 K, SiVs experience temperature-dependent spin decoherence (20, 21) primarily through a resonant two-phonon process that involves the orbital upper-branch ground states as an intermediate state. Electrons from the qubit manifold are excited by thermal phonons to the upper orbital states for a random amount of time before decaying via another phonon back to the lower orbital states. In this process, the electron spin state is preserved as our magnetic fields are aligned with the SiV symmetry axis, but the different precession rate in the upper branch means that the electron obtains a random phase and becomes dephased after many such stochastic excitations. The rate of this two-phonon process is proportional to the Bose-Einstein distribution of phonon occupancy at the ground state splitting energy Δ_{GS} scaling as $(e^{h\Delta_{GS}/k_B T} - 1)^{-1}$, so we expect that SiVs with a larger Δ_{GS} will have coherence times that are less sensitive to increasing temperature (16, 20). There is additionally a weaker decoherence effect from the direct single-phonon spin-flipping process that limits electron T_2 through electron depolarization, but the rate of this process is lower as it scales with the relatively smaller qubit frequency $\omega_{\text{qubit}} \sim 12$ GHz as ω_{qubit}^3 (20). The dominant limitation on the value of T_2 for our SiV electron spins is therefore two-phonon processes to the upper ground state branches.

To verify that the sustained long coherence times of our device at high temperatures (Fig. 3A) is indeed due to its large Δ_{GS} suppressing the rate of two-phonon events, we can look at a closely related quantity, the electron spin relaxation rate T_1^{-1} . The spin lifetime is also limited

by the two-phonon process (16, 20) since the electron has a small chance of undergoing a spin-flipping decay from the upper orbital state when the magnetic field is not perfectly aligned (39); this is also known as the Orbach process (40). The direct single-phonon spin-flipping process also has a contribution to T_1^{-1} that is relevant at lower temperatures ($k_B T / \hbar \ll \omega_{\text{qubit}}$), while the two-phonon process dominates at higher temperatures comparable to the ground state splitting ($k_B T / \hbar \sim \Delta_{GS}$) (16, 20). The crossover point between these two mechanisms with increasing temperature will be observable as an inflection with a sharp decrease in spin lifetime and coherence.

Our results for the temperature dependence of T_1^{-1} for several SiVs in the same device are shown in Fig. S12. These SiVs have a range of ground state splittings Δ_{GS} arising from the inhomogeneous distribution of local strain within the same nanocavity. We normalize the T_1 for each SiV to their value at the base temperature of 100 mK ($T_{1,0}$) since the absolute value of T_1 for each SiV can differ due to how well aligned the magnetic field is to their symmetry axes, but we are only interested in the scaling of T_1 with temperature. The T_1 behavior is fitted with the single-phonon and two-phonon process scaling behaviors as described in (20).

As expected, the SiVs with larger ground state splitting, including the main SiV used in this work ($\Delta_{GS} = 554$ GHz), exhibit the smallest T_1 drop with increasing temperature in this temperature range. Furthermore, the temperature at which two-phonon processes begins to dominate – as indicated by the kink in the slope that precedes a rapid decrease in T_1 – increases with increasing Δ_{GS} as predicted from the theory. In particular, we find another device with $\Delta_{GS} = 416$ GHz that is also highly strained and has excellent T_1 behavior up to 1.5 K, further confirming that the large Δ_{GS} from strain is a key determinant towards the high-temperature operation of our current device via the suppression of two-phonon processes.

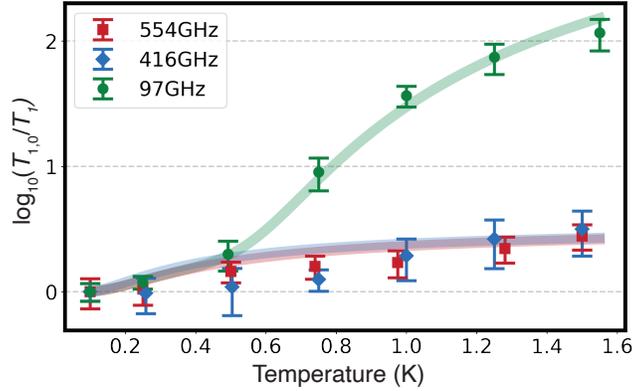


Fig. S12. Comparison of electron T_1^{-1} (normalized to $T_{1,0}$, their value at 100 mK) as a function of temperature for various SiVs in the same device with different ground state splittings Δ_{GS} . The ground state splittings Δ_{GS} for all SiVs were measured at zero magnetic field at 100 mK by driving resonantly on the spin-conserving C transition and collecting fluorescence from the spin-flipping D transition (41). The device used in this work has $\Delta_{GS} = 554$ GHz.

11 Distribution of SiV Strain Across Devices

We characterize the approximate ground state splitting distribution in our devices to estimate the feasibility of preselecting SiVs in nanocavities for high strain. First, we perform direct measurements of the ground state splitting with resonant excitation for several SiVs that were found in the same device as the main SiV used in this paper (Fig. S13A). We send light resonant with an individual SiV’s C transition and measure the outgoing light from the cavity with a spectrometer at zero magnetic field. The measured spectrum shows two peaks corresponding to the SiV C and D transitions, and the difference in frequency between these peaks is the ground state splitting value for the given SiV. We measure 9 different SiVs in our device and find 2 highly strained SiVs with $\Delta_{GS} > 400$ GHz.

For a more statistically significant sample of the strain distribution, we perform a large-scale spectroscopy of Δ_{GS} on a separate chip with 12 cavities that was fabricated in an identical manner. The measurement is taken in a 4 K closed-cycle helium cryostat (Montana Instruments) with no magnetic field applied. The SiVs are illuminated off-resonantly at 520 nm through a free-space port, after which fluorescence counts are collected confocally and sent to a spectrometer (Fig. S13C). The observed peaks are then grouped into sets of 4 peaks corresponding to the four A – D transitions of single SiVs fulfilling the requirements $\omega_A - \omega_B = \omega_C - \omega_D > 48$ GHz (minimum ground state splitting for an unstrained SiV) and $\omega_A - \omega_C = \omega_B - \omega_D > 259$ GHz (minimum excited state splitting for an unstrained SiV) (42).

The number of SiVs per cavity is not known with certainty, but for a given cavity we can find an estimated minimum number (so that all peaks that can possibly be matched with at least 1 SiV are assigned a match) and a maximum number of SiVs (all allowed sets of 4 peaks

are assigned to a separate SiV). For a given choice of the mean number μ of SiVs per cavity, we can obtain a sample of the estimated Δ_{GS} distribution by first sampling a value n for the number of SiVs per cavity from the normal distribution $n \sim \mathcal{N}(\mu, \sigma = 2)$ (clipped within the bounds explained above), and then sampling a possible assignment configuration of n SiVs to the peaks in the photoluminescence spectra, from which their Δ_{GS} may be extracted. This sampling procedure is repeated 10,000 times for each choice of the mean number of SiVs per cavity.

Fig. S13B shows the resulting ground state splitting distributions for the minimum (8.8) and maximum (98.8) number of SiVs assigned per cavity during the fitting procedure. We also show the results for 22 SiVs per cavity, which is an estimate of the actual value based on resonant excitation measurements performed on the chip used in this experiment. We find that the proportion of SiVs with large ground state splitting ($\Delta_{GS} > 400$ GHz) is $12.2 \pm 1.4\%$ for 22 SiVs per cavity, and the lowest estimate of this proportion is $11.0 \pm 1.1\%$ when assigning the maximum number of SiVs per cavity.

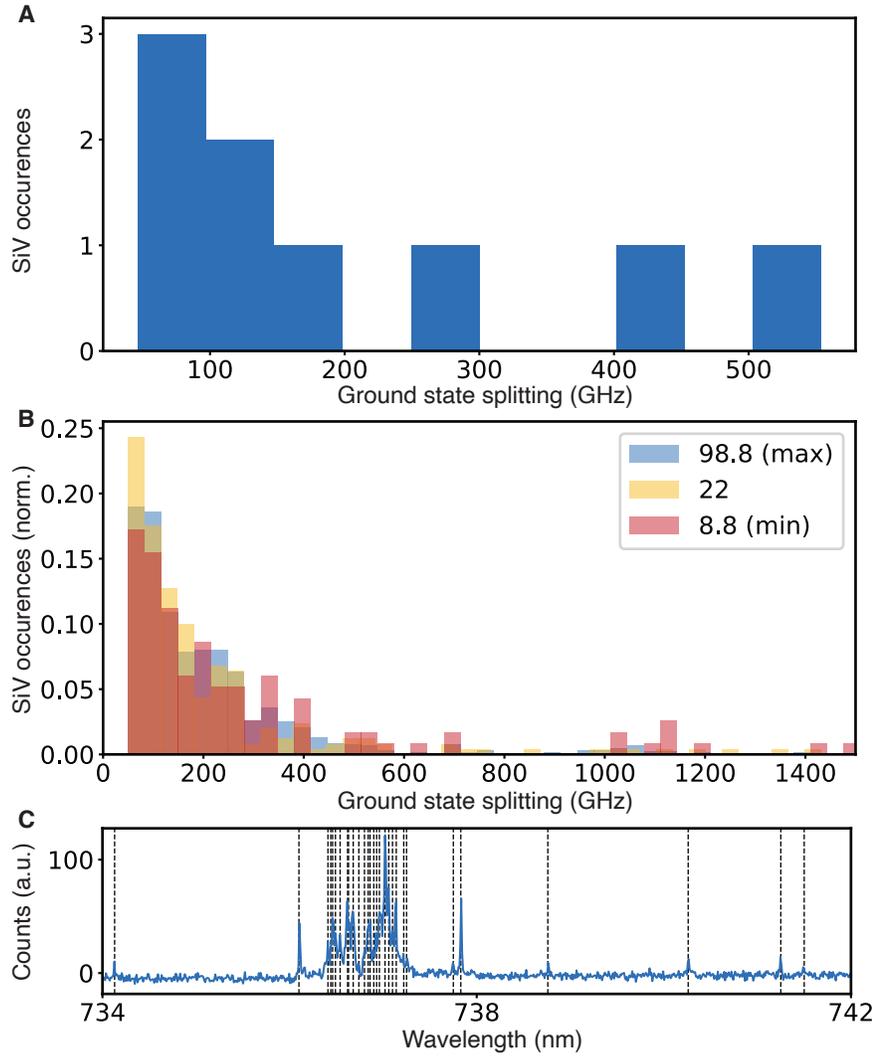


Fig. S13. (A) Distribution of Δ_{GS} measured in the same device as the main SiV used in this paper using resonant excitation. (B) Distribution of Δ_{GS} in a different device using off-resonant excitation. The different histograms correspond to the different numbers of average SiVs per cavity used during the fitting procedure. (C) Photoluminescence spectrum of one of the measured cavities with matched peaks marked with vertical lines.

12 Spin-fluctuator model and motional averaging

The temperature dependence of the coherence times for the electron and nucleus was investigated in the main text. Here, we present the models of surrounding spin and phonon environment that were used to fit the observed data. The electron coherence time is primarily determined by coupling to a non-Markovian magnetic noise bath and dephasing from phonon excitation processes. The phonon excitation processes we include are a resonant 1-phonon process and a resonant 2-phonon process with rate scalings taken from (20). The depolarization timescale $T_{1,e}$ is much larger than any measured $T_{2,e}$, so we assume it does not significantly contribute to dephasing. The resulting temperature dependence of the electron coherence time that we used for the fitting is:

$$T_{2,e}^{\text{tot}}(T) = \left(T_{1\text{-phonon}}^{-1}(T) + T_{2\text{-phonon}}^{-1}(T) + T_{\text{noisebath}}^{-1} \right)^{-1}$$

On the other hand, for the dependence of the nuclear coherence time on temperature, we incorporate the contributions from three separate decoherence processes. First, depolarization of the electron causes dephasing of the nucleus due to the strong hyperfine interaction between the electron and the nucleus resulting in differential phase accumulation. The rate of this process can be measured through the electron $T_{1,e}$ (Fig. S12). Second, we expect the nuclear coherence to be affected by a bath of ^{13}C nuclear spins. As the ^{13}C nuclear spin couples very weakly to phonons (43), this spin bath is modeled to be temperature-independent. Finally, we add another spin bath that is modelled as a bath of two-level spins with phonon-addressable transitions. We model this bath as causing discrete frequency fluctuations of the nuclear spin with a Monte Carlo simulation (44), with the rate of these fluctuations scaling with the temperature-dependent phonon density at the characteristic transition frequencies of the bath (20). Through the effects of motional averaging (45, 46), this last bath leads to the observed increase in nuclear coherence with temperature. This is because as the temperature increases, so does the rate of fluctuations of the phonon-sensitive spin bath; when this fluctuation timescale becomes much smaller the decoupling sequence timescale, the fluctuations average out and cause the nucleus to feel a weaker averaged perturbation from the bath. This leads to the coherence time increase as seen in Fig. S14. Combining all processes, the resulting temperature dependence of the nuclear coherence time is:

$$T_{2,\text{tot}}(T) = \left(T_{1,e}^{-1}(T) + T_{2,^{13}\text{C}}^{-1} + \sum_i T_{2,\text{ph}}^{-1}(T, \omega_i) \right)^{-1}$$

where $T_{1,e}(T)$ is the electron coherence time, $T_{2,^{13}\text{C}}$ is the temperature-independent dephasing from ^{13}C nuclear spins, and $T_{2,\text{ph}}(T, \omega_i)$ is dephasing from a spin bath of two-level systems with a phonon-sensitive transition at frequency ω_i . We fit the temperature coherence data in Fig. S14 and find two characteristic phonon-addressable transition frequencies at 26 GHz and 34 GHz, which could potentially correspond to reconfigurations of microstrain in the diamond lattice, or

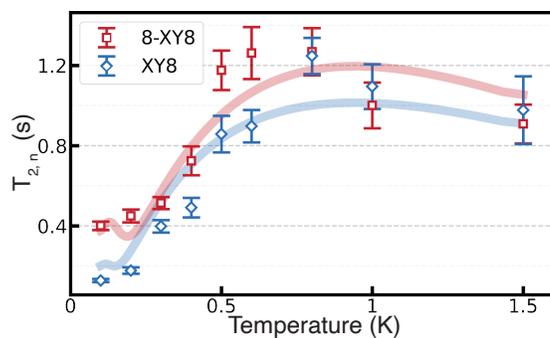


Fig. S14. Coherence time of the nuclear spin as a function of temperature. Solid line is a fit to the decoherence model of the spin environment.

orbital state splittings of nearby defects.

13 Optimization of the phase readout

Many quantum networking protocols, including entanglement distillation and swapping with nondeterministic heralded entanglement schemes, require multiple readouts of the electron state in the middle of the experimental sequence while preserving the state of the nucleus (47). To maximize the number of readouts and their fidelity, we can optimize the cavity-QED parameters of the SiV in the nanophotonic cavity together with the readout laser parameters.

A useful optimization metric as a function of laser frequency f is the ratio of the trace distances of electron states over the trace distances of the nuclear states since the trace distances are proportional to the maximum distinguishability between the respective spin states. We only consider the reflected photonic states for the electron since this is the only signal we can collect for state readout. On the other hand, we calculate the overlap over all three paths: reflection, scattering, and transmission for the nuclear trace distance. The key is to maximize the electron state information and minimize the nuclear state disturbance by each measurement (48).

The measurement constitutes sending in a coherent state to the cavity as an input and getting spin-state dependent output in three ports reflection, transmission and scattering:

Before interaction:

$$|\psi_0\rangle = |\alpha\rangle \otimes (|\uparrow\rangle + |\downarrow\rangle)$$

After interaction:

$$|\psi_1\rangle = |R_{\uparrow}\alpha\rangle \otimes |\uparrow\rangle + |R_{\downarrow}\alpha\rangle \otimes |\downarrow\rangle$$

where we can only collect the signal from the reflected port. The output coherent states are pure states that encode the information about the spins in our system. Since no two coherent states are orthogonal, the trace distance between the output states are not zero: $d_{\uparrow\downarrow_e} = \sqrt{1 - \langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle_e^2} \neq 0$. This results in an upper bound of maximum probability to distinguish two electron spin states using the reflected coherent states:

$$d_{\uparrow\downarrow_e} = \sqrt{1 - \langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle_e^2}$$

The overlap $\langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle$ determines how well we can distinguish the spin states in our measurement. The farther apart the spin-dependent coherent states are in phase space, the more distinguishable the spin states are in the measurement. There are two ways to change the phase space distances: the first one is to pick a frequency at which the two cavity reflection coefficients are farthest from each other on the circle in phase space traced out by the complex reflection coefficients as the laser frequency is varied (Fig. S15C), and the second one is to increase the amplitude $|\alpha|$, which scales the diameter of the circle by shining more photons onto the cavity.

We need to include all output ports for the nuclear trace distance since decoherence is not

dependent on our efficiency in collecting data from multiple ports.

$$d_{\uparrow\downarrow n} = \sqrt{1 - \langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle_n^2 \langle T_{\downarrow}\alpha | T_{\uparrow}\alpha \rangle_n^2 \langle S_{\downarrow}\alpha | S_{\uparrow}\alpha \rangle_n^2}$$

where $d_{\uparrow\downarrow n}$ is the trace distance between the reflected photonic states for the two nuclear states. Finally, our optimization function is the ratio of the probabilities to distinguish the electron state over nuclear states as a function of the laser frequency f :

$$\begin{aligned} \text{Optimizer}(f) &= \frac{d_{\uparrow\downarrow e}}{d_{\uparrow\downarrow n}} \\ &= \frac{\sqrt{1 - \langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle_e^2}}{\sqrt{1 - \langle R_{\downarrow}\alpha | R_{\uparrow}\alpha \rangle_n^2 \langle T_{\downarrow}\alpha | T_{\uparrow}\alpha \rangle_n^2 \langle S_{\downarrow}\alpha | S_{\uparrow}\alpha \rangle_n^2}} \end{aligned}$$

Note that the reflection coefficients for the electron spin-down and spin-up states also depend on the nuclear state, so the numerator implicitly includes an average over both nuclear states. The denominator similarly includes an implicit average over both electron states.

For any given choice of cavity-QED and SiV parameters, we can minimize this optimizer to find an optimum choice of laser frequency for electron state readout while preserving nuclear coherence. We can then compute the number of photons required at this frequency to perform an electron state phase-based readout at 95% fidelity, which gives us the nuclear decoherence per readout. This lets us compute the number of electron readouts possible before we experience a $1/e$ nuclear state decoherence, which we then optimize across all practical device and SiV parameters.

We find that systems where more electron readouts are possible for a given amount of nuclear decoherence typically: (I) have a larger g and have over-coupled cavities. These parameters maximize the amount of incoming light directed into the reflected port, thus increasing the amount of information gained per photon sent; (II) have a large electron splitting Δ_e while keeping the nuclear state splitting Δ_n constant, which can be achieved by ramping up the magnetic field. However, operation at the correspondingly higher microwave frequencies have limiting factors, including signal delivery inefficiencies and an eventual reduction of $T_{1,e}$ due to an increased phononic density of states. In particular, for achievable parameters of $\{g = 2\pi \times 7, \kappa_{in} = 2\pi \times 180, \kappa_{tot} = 2\pi \times 250, \Delta_a$ (cavity detuning) $= 100, \Delta_e = 1.5, \Delta_n = 0.05, \gamma = 0.16\}$ GHz, we expect to be able to do approximately 90 electron readouts at 95% fidelity without postselection before we experience $1/e$ nuclear decoherence.

The optimum frequency above gives the best probe point if one is able to detect both amplitude and phase differences in the reflected signal. Practically, it may be more convenient to sacrifice some electron state distinguishability or nuclear state fidelity by purely optimizing for the amplitude contrast or phase difference between the reflected probe coherent states. This is especially useful for designing “one-shot” state readouts similar to our resonant readout or a homodyne readout.

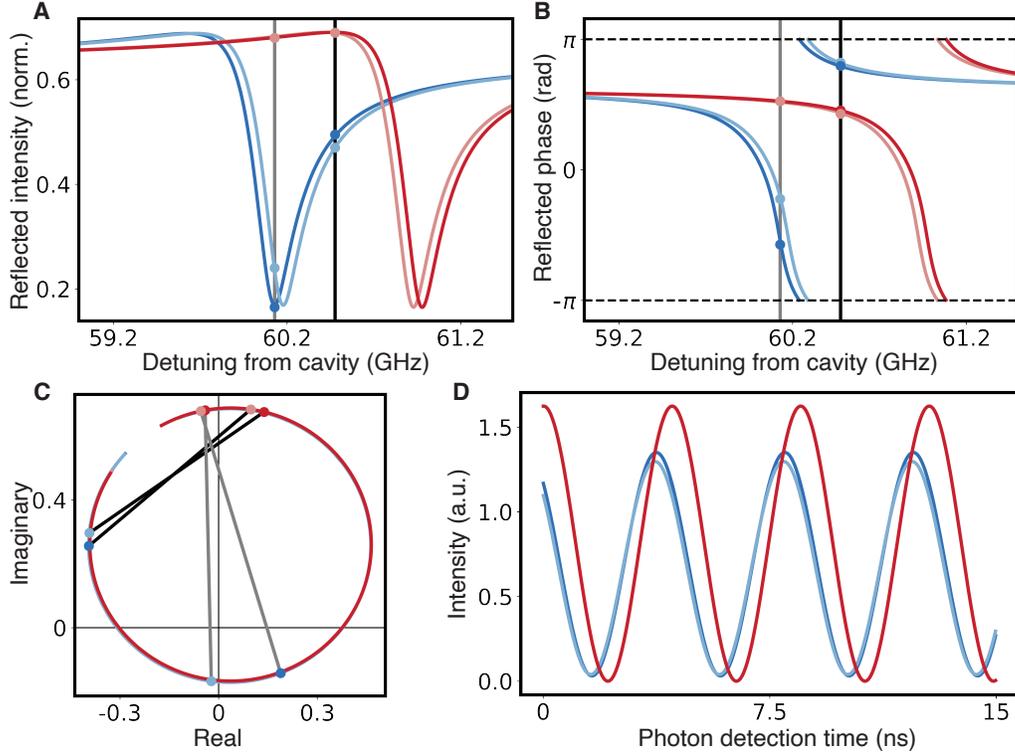


Fig. S15. Cavity reflection spectra for two-qubit states: $|\downarrow_e \downarrow_n\rangle$, $|\downarrow_e \uparrow_n\rangle$, $|\uparrow_e \uparrow_n\rangle$ and $|\uparrow_e \downarrow_n\rangle$ in dark/light blue and dark/light red. Cavity-QED parameters was chosen for a typical device with $\{g = 2\pi \times 4, \kappa_{in} = 2\pi \times 190, \kappa_{tot} = 2\pi \times 250, \Delta_a = 60, \Delta_e = 0.8, \Delta_n = 0.05, \gamma = 0.16\}$ GHz. **(A, B)** Magnitude and phase of a reflected photon. Gray line: Frequency of maximum amplitude difference between electron states, used for resonant readout. Black line: Optimized frequency with maximum phase-space separation for the electron states and minimum nuclear decoherence, used for phase readout. **(C)** Complex reflection amplitude for the four different states as the laser frequency is varied, which forms a circle in the complex plane. Points indicate the resonant readout frequency (gray) and optimized nuclear state-preserving readout frequency (black). **(D)** Phase-measurement signal where the optimum readout frequency for the given cavity-QED parameters is interfered with another readout frequency far from SiV resonance.