Supplementary material for: Engineering a diamond spin-qubit with a nano-electro-mechanical system

Young-Ik Sohn*[1], Srujan Meesala*[1], Benjamin Pingault*[2], Haig A. Atikian[1], Jeffrey Holzgrafe[1,2], Mustafa Gündoğan[2], Camille Stavrakas[2], Megan J. Stanley[2], Alp Sipahigil[3], Joonhee Choi[1,3], Mian Zhang[1], Jose L. Pacheco[4], John Abraham[4], Edward Bielejec[4], Mikhail D. Lukin[3], Mete Atatüre[2], and Marko Lončar[1]

[1]John A. Paulson School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA
[2]Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, UK
[3]Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA
[4]Sandia National Laboratories, Albuquerque, NM 87185, USA

1 Fabrication

The diamond NEMS (nano-electro-mechanical system) device used in this work was fabricated in three steps in the following order: (i) fabrication of bare diamond cantilevers, (ii) creation of silicon vacancy color centers, and (iii) deposition of electrodes. We use commercially available, ⟨100⟩-cut, ultra-high purity, single-crystal diamond (type IIa, nitrogen concentration less than 5 ppb) synthesized by chemical vapor deposition (CVD) [1].

Cantilevers are fabricated in two steps. First, diamond with patterned electron-beam resist is etched vertically with oxygen plasma [2]. These vertically-etched structures are then made free-standing by etching the sample at a tilted angle. Specifically, we employ an oxygen-plasma assisted ion-milling process in which the sample is mounted at an angle that is manually adjustable within a few degrees of precision. An illustration of this process, and the resulting suspended structure is schematically shown in Fig. S1(a).The etching occurs over a period of a few hours, during which the stage is rotated constantly. Further discussion

*These authors contributed equally

of these techniques can be found in [3, 4].

After cantilever fabrication, silicon ions (Si^+) are implanted at target spots on the cantilevers using a custom focused-ion-beam (FIB) system at Sandia National Labs. The spot size of the ion-beam on the sample is 40 nm, and is expected to determine the lateral precision of the implantation procedure. The beam energy is chosen to be 75 keV, which is predicted to yield a mean implantation depth of 50 nm with a straggle of 10 nm according to Stopping and Range of Ions in Matter simulations. Further details of the FIB implantation procedure can be found in [5, 6]. After FIB implantation, the sample is subjected to a tri-acid clean (1:1:1 sulfuric, perchloric, and nitric acids), and a three-step high-temperature high-vacuum annealing procedure [7, 8] in an alumina tube furnace. The annealing sequence followed comprises steps at 400°C (1.5°C per minute ramp, 8 hour dwell time), 800°C (0.5°C per minute ramp, 12 hour dwell time), and 1100°C (0.5°C per minute ramp, 2 hour dwell time). During the entire procedure, the pressure is maintained below 5×10^{-7} torr. Annealing generates a small amount of graphite on the diamond surface, which is subsequently etched away by a tri-acid clean. Following this step, we perform a cleaning in piranha solution to ensure a high level of oxygen-termination at the diamond surface. With regards to conversion efficiency, we implant approximately 50 Si^+ ions per target spot on the sample, and typically generate 1-3 SiVs at each spot after annealing.

Subsequently, electrode patterns are made by a conventional bi-layer PMMA process followed by metal evaporation. Since the distance between the top surface of the cantilever and the bottom substrate is approximately 4 μm, bi-layer PMMA is spun multiple times until the cantilevers are buried completely. Patterns are written by electron-beam lithography, and metals are evaporated to define the electrodes. Detailed fabrication steps are schematically shown in Fig. S1(b). Here, the triangle represents the cantilever, and the pedestal to the right of the triangle is the location of the bonding pad for electrical contact. The bi-layer PMMA process is repeated twice - first, to define the bonding pad, and second, to define the electrode pattern near the cantilever. This is because, we use a 200 nm thick gold layer for the bonding pad, but only a 10 nm thick tantalum (Ta) layer for the cantilever electrodes. Fig. S1(c) shows the scheme to connect electrodes on top of the cantilevers to the bonding pad on the diamond pedestal. Electrodes on the substrate below the cantilevers are connected to a second bonding pad (now shown in the figure) that is directly on the surface of the diamond.

We now discuss the choice of 10 nm tantalum film for our cantilever electrodes. For five different metals we have tested as a cantilever electrode material (aluminium, chromium, copper, titanium and tantalum), our device always shows a continuous, non-zero leakage-current upon applying voltage. While the exact reasons for this leakage-current on our sample, in particular, are unknown, there have been numerous studies on the surface conductivity of diamond under various conditions [9]. For aluminium, chromium, copper and titanium, this leakage current destroys the electrode when a high voltage (in the few hundred volts range) is applied. The destruction of electrodes appeared to be the result of melting or bursting of the thin metal film, likely caused by Joule heating [10]. Tantalum is one of the metals with the highest melting and evaporation points among those available for e-beam evaporation. We find that devices with tantalum electrodes are robust enough to operate at very high

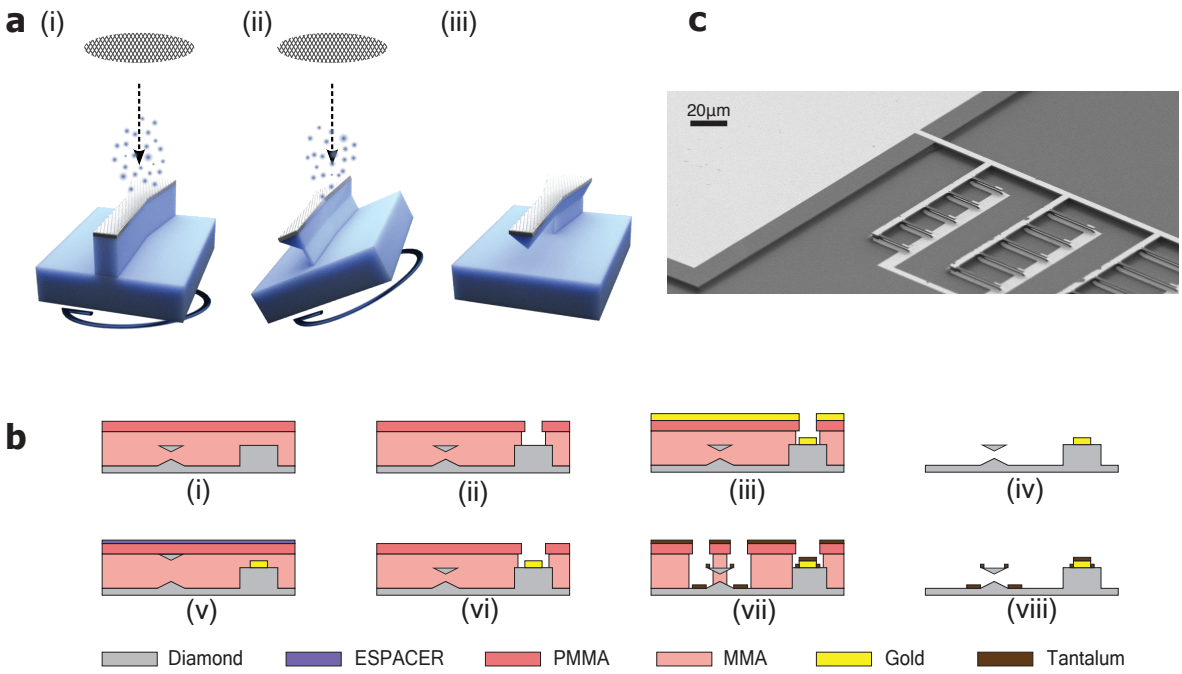

Figure S1: (a) Schematic of oxygen-plasma assisted ion-milling process for angled-etching of diamond cantilevers. The ion beam is directed at the diamond sample, with a vertically-etched device pattern. The tilted stage is continuously rotated during the etching process. After the cantilevers are freely standing, the etch-mask is stripped. (b) Fabrication process for the placement of electrodes. First, the coarsely aligned bonding pad is defined with a bi-layer PMMA process followed by gold evaporation. Then the same process is repeated to define tantalum electrodes near cantilevers, but with better alignment precision. Conductive layer (ESPACER 300Z) on top of the cantilever is helpful for precise alignment. (c) SEM image of the complete chip showing connection between the bonding pad and electrodes on top of the cantilevers.

applied voltage (∼600 V across an electrode gap of approximately 4 μm, which corresponds to an electric-field of 1.5 MV/cm). The thickness of tantalum is kept below 10 nm in order to avoid thin film-induced stress in the cantilever, which leads to pre-strained SiV centers.

2 Device design

An important figure of merit for our NEMS device is the maximum achievable strain at the location of SiV. In this section, we discuss two key design aspects that need to be considered towards this goal: (i) 'pull-in instability', and (ii) practical limits for high voltage operation.

Pull-in instability is a well-known phenomenon for an electrostatic actuator made with a parallel plate capacitor as shown in Fig. S2(a). In these devices, voltage is applied to induce an electrostatic force between two plates, where either one or both of them are free to move. Upon applying a voltage, the capacitor deforms until it reaches equilibrium, when there is a

balance between the electrostatic force, and the restoring force exerted by the elasticity of the material. The net force acting on the free top plate in Fig. S2(a) can be modeled as

$$F(x,V) = -\frac{\partial U(x,V)}{\partial x} = \frac{1}{2}\varepsilon A\frac{V^2}{(d-x)^2} - kx \tag{1}$$

where x is the displacement of the plate, $U(x,V)$ is the potential energy and ε is the permittivity of the material between the two plates. A is the area of the capacitor, V is the voltage applied, d is the distance between the two plates at 0 V, and k is the spring constant, respectively. By integrating equation (1), we can calculate $U(x,V)$ at various voltages as shown in Fig. S2(b). The local minimum in the potential represents a condition of stable equilibrium. As the voltage is increased, the local minimum shifts towards the bottom plate, indicating that the top plate gets displaced downwards, thereby reducing the capacitor gap. When the voltage changes from $3V_0$ to $4V_0$ in the Fig. S2(b), the stable local minimum disappears. This occurs, when the top plate is displaced by about one-third of the initial gap between the two plates, i.e. when $x = d/3$. At this point, the system reaches a condition in which the top plate snaps down to the bottom plate. Our device is a slight variation of this conceptual model, and hence, the maximum deflection of our cantilever will be limited by pull-in instability (but not at exactly $x = d/3$).

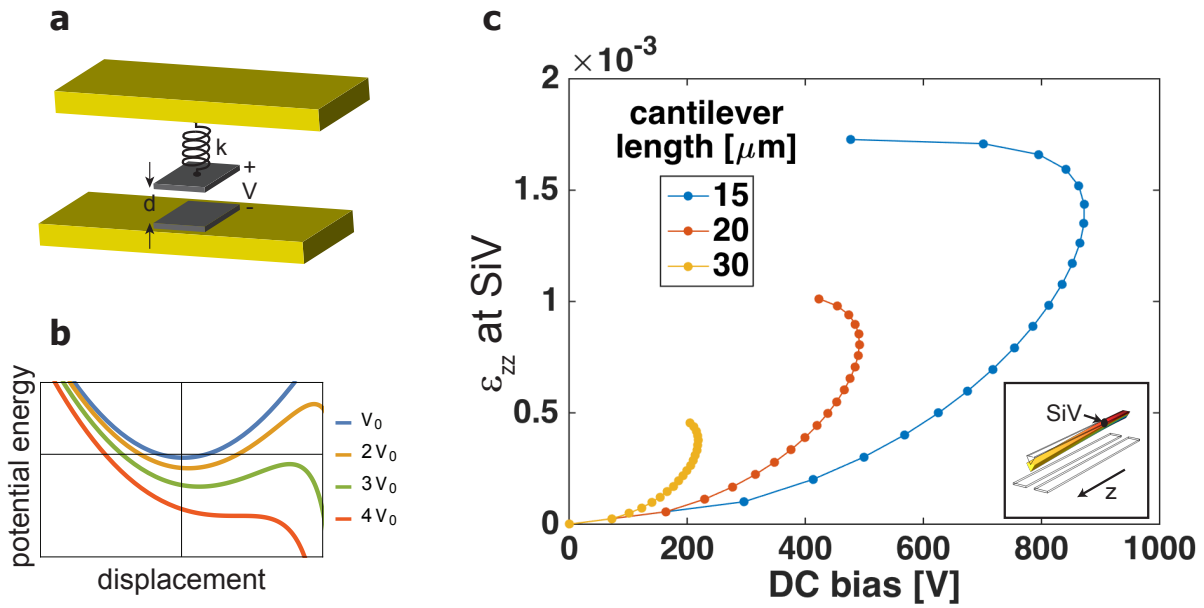

Figure S2: (a) Illustration of a parallel plate capacitor with one freely movable plate (top), and one fixed plate (bottom). The top electrode can be actuated by applying a voltage. (b) Potential energy of the system in (a) with the different voltages. The stable minimum in the potential disappears, when the system reaches the condition of pull-in instability at a voltage of $4V_o$. (c) FEM simulation of the strain-component along the long-axis of the cantilever (most dominant strain-tensor component) near the clamp of the cantilever (inset). Turnaround points in the graph represent pull-in instabilities.

A 3D finite element method (FEM) calculation can be used to simulate pull-in instability accurately for complex structures [11]. Typically, a simulation can be run by setting a particular voltage on the electrode, and solving for the resultant deformation of the structure, thereby arriving at the strain profile inside the cantilever. One can also run the inverse of this procedure. By setting a target displacement of the cantilever tip, the voltage required to achieve this displacement can be calculated. Such an inverse calculation can help arrive at the condition for pull-in instability. Fig. S2(c) shows the results of a simulation run so as to solve this inverse problem. Strain at the SiV location is plotted as a function of the voltage for different cantilever lengths. Turnaround points represent the pull-in instability condition at which both the displacement of the cantilever-tip and the applied voltage reach the maximum value possible before the cantilever snaps down. Fig. S2(c) provides two important conclusions: First, for a given voltage, longer cantilevers provide larger strain for a given voltage, because they have a smaller spring constant. Second, the maximum attainable strain is higher for shorter cantilevers, because they reach the pull-in instability condition at a higher voltage. Therefore, shorter devices are preferred to generate high strain, when arbitrarily high voltage can be applied.

In practice, however, there are mechanisms that limit the maximum possible voltage e.g. Townsend breakdown, field emission and surface current [12, 13, 14], all of which can be significant depending on experimental conditions. With the fabrication method described in Section 1, our devices could be operated safely up to voltage as high as 600 V under high vacuum ($\sim 10^{-7}$ torr) at cryogenic temperature (4 K). Given that the minimum electrode gap is 4 μm, this condition corresponds to an electric field of approximately 1.5 MV/cm. Experiments described in Section 10 are carried out in a helium closed-cycle cryostat with the sample surrounded by helium exchange gas at a pressure of 1 mbar. Under these conditions, we observed safe operation up to 500 V. The maximum voltage in this setup is thought to be limited by dielectric breakdown of helium gas.

Considering all the design limitations discussed above, we chose cantilevers of width 1.2-1.3 μm and length 25-30 μm for the experiments in this work.

3 Cantilever-induced strain in SiV frame

The diamond samples used in our study are [001] cut on the top surface, and the direction of the cantilever long-axis is [110]. In general, there are four possible equivalent orientations of SiVs - [111], [$\bar{1}\bar{1}$1], [1$\bar{1}$1], [$\bar{1}$11] - in a diamond crystal. These are indicated by the four arrows on top of the cantilever in Fig. S3(a). For each SiV-class, this arrow represents the highest symmetry axis or the Z-axis in its internal frame (see inset of Fig. S3(a)). Now, we note that the dominant strain-component applied by the cantilever is uniaxial, and directed along the cantilever long-axis, [110] as shown in Fig. 1(c) in the main text. Upon deflection of the cantilever, application of strain along this direction breaks the equivalence of the four SiV orientations, and leads to two classes indicated by the blue and red colored arrows in Fig. S3(a). By symmetry, within each class, we expect the same response to strain. The blue SiVs, which comprises [1$\bar{1}$1], [$\bar{1}$11] SiVs, is oriented perpendicular to the cantilever

long-axis, [110], which happens to be the Y-axis in the internal frame of these SiVs (see inset of Fig. S3(a)). Thus we expect predominantly uniaxial strain along the internal Y-axis of these SiVs. On the other hand, the red class, which comprises [111], $[\bar{1}\bar{1}1]$ SiVs is not orthogonal to the cantilever long-axis. The strain-tensor of these SiVs in their internal frame is non-trivial, but unlike the blue SiVs, they are expected to experience a significant strain component along the internal Z-axis.

To estimate strain at the SiV location, when a voltage is applied to the device, we run an FEM simulation with a geometry close to that of the fabricated device (dimensions are measured from SEM images). The simulation gives the strain tensor, a 3-by-3 matrix, at any location inside the cantilever. The resultant strain tensor is in the reference frame of the device, and can subsequently be rotated into the internal reference frame of SiV to gain more insight. For an SiV which is at a distance of 2.5 μm from the cantilever-clamp, and 50 nm below the surface, the simulated strain tensor was rotated into the internal reference frame for both red and blue classes, and plotted in Fig. S3(b) and (c), respectively. The magnitude of each strain tensor component is normalized to the maximum value among all of its components. For the red class, ϵ_{ZZ} is the largest value among all nine components, and for the blue class, ϵ_{YY} is the largest component. For this reason, we conveniently refer to the red and blue classes as 'axial-orientation' and 'transverse-orientation' SiVs, respectively. This nomenclature is used with the caveat that it is specific to the situation of predominantly [110] uniaxial strain applied with our cantilevers. In general, as we will see in Section 6.1, one needs knowledge of the full strain-tensor to describe the response of the electronic levels of the SiV. For instance, for the axial-orientation SiV in Fig. S3(b), we see that the shear component ϵ_{YZ} is quite significant compared to the axial component, ϵ_{ZZ}.

In this manner, the strain-tensor at the SiV site is simulated as a function of applied voltage, and is used to fit our experimental data to the theoretical strain-response model in Section 6.2. The final results are shown in Fig. 2(b) of the main text for a transverse-orientation SiV, and in Fig. S6 for an axial-orientation SiV, respectively.

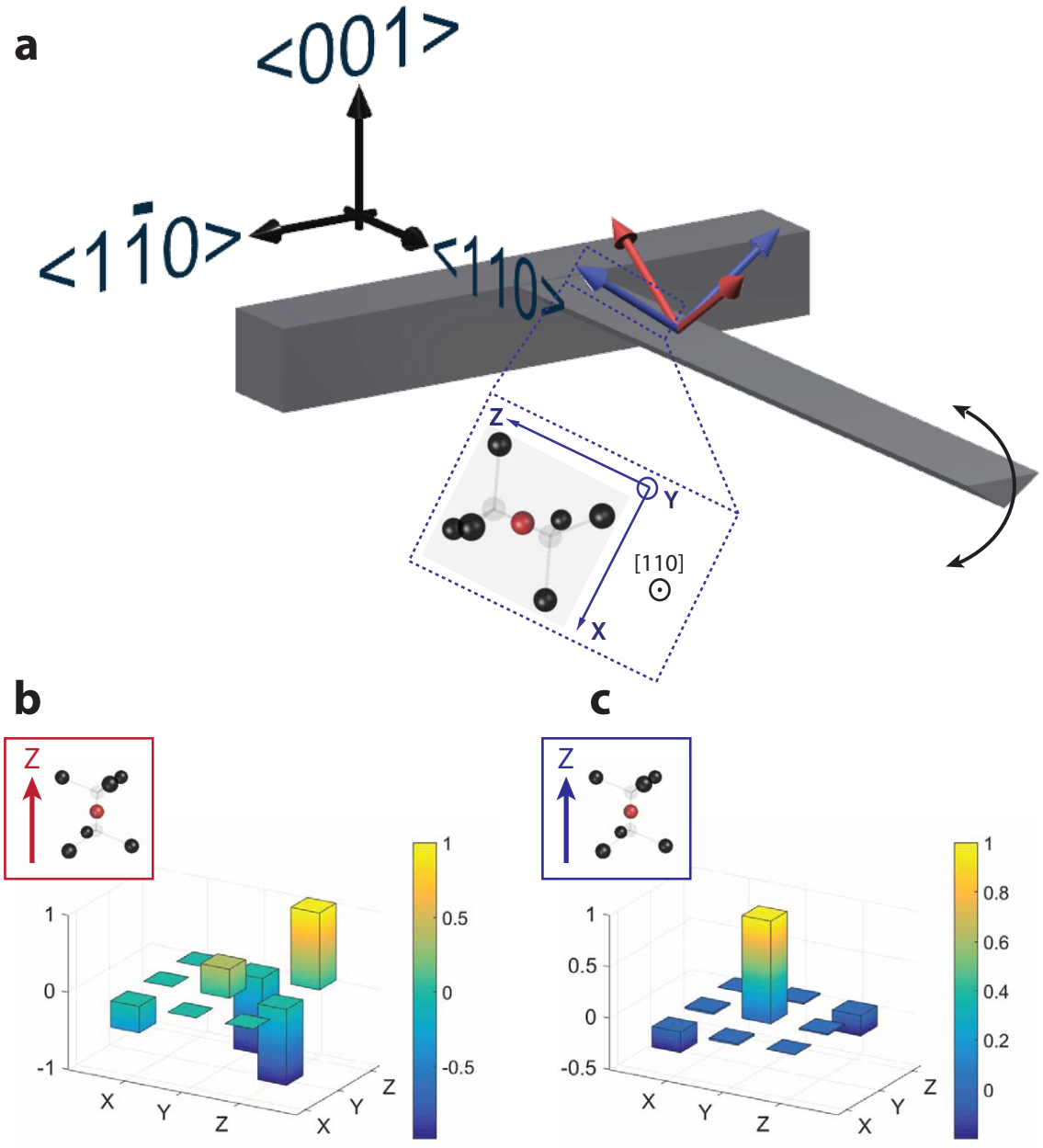

Figure S3: (a) Orientation of cantilevers and SiVs with respect to diamond crystal axes. Four possible orientations of the highest symmetry axis of an SiV are indicated by the four arrows above the cantilever. Under application of strain, these orientations can be grouped into two classes indicated by red and blue colored arrows. Inset shows the molecular structure of a blue SiV with its internal co-ordinate frame as viewed in the plane normal to the [110] direction. The Z-axis is the highest symmetry axis, which defines the orientation of the SiV. For blue SiVs, the Y-axis coincides with the long-axis of the cantilever, [110]. (b) Normalized strain-tensor components for a red SiV. This is an axial-orientation SiV as referred to in the text. (c) Normalized strain-tensor components for a blue SiV. This is a transverse-orientation SiV as referred to in the text.

4 Strain-dependent photoluminescence measurements

4.1 Experimental setup

The sample is cooled down to a nominal temperature of 6 K inside a Janis ST-500 continuous helium-flow cryostat. The cryostat is mounted under a home-built scanning confocal microscope with a 0.9 NA 100×, 1 mm working distance objective (Olympus MPLFLN 100X) housed inside the cryostat. SiV centers are identified via non-resonant excitation with a 703 nm laser diode (Thorlabs LP705-SF15), and collection of zero-phonon-line (ZPL) fluorescence in a narrow bandwidth of 10 nm around 737 nm. For resonant excitation of ZPL transitions, we use a tunable continuous-wave Ti-sapphire laser (M-Squared Solstis), and collect the resulting fluorescence in the phonon-sideband (above 750 nm). Mode-hop-free tuning of the laser is achieved over the scan range of interest by using feedback from a wavemeter (High Finesse WS7). A CW 532 nm laser is periodically pulsed using an acousto-optic modulator (Crystal Technology 3080) to maintain the negative charge state of SiV(-) centers [5]. Collected fluorescence (gated off during green excitation) is sent to an avalanche-photodiode to measure the photon-count rate. DC voltage for cantilever-deflection is supplied from a Stanford Research Systems PS300 high-voltage source. As an added precautionary measure, the weak leakage-current in the circuit (typically below 100 nA) discussed in Section 2 is monitored via a Keithley 2400 source-meter. Strain-response measurements involve taking resonant excitation spectra as the voltage applied to the device is steadily increased.

4.2 Strain-dependent resonant excitation (PLE) spectra

Two distinct types of response in optical transitions are observed in strain-tuning measurements (Figs. S4, S5). These are correlated with the two classes of SiV orientations discussed in Section 3. The orientation of the SiVs measured in our experiments is inferred from the polarization-dependence of their optical transitions at zero strain. In particular, Fig. 2(b) of the main text depicting strain-tuning of a transverse-orientation SiV shows the peak wavelengths obtained from Fig. S4 with strain values estimated using FEM simulation as described in Section 3. Similarly, strain-dependent spectra are taken for an axial-orientation SiV, and are plotted in Fig. S5.

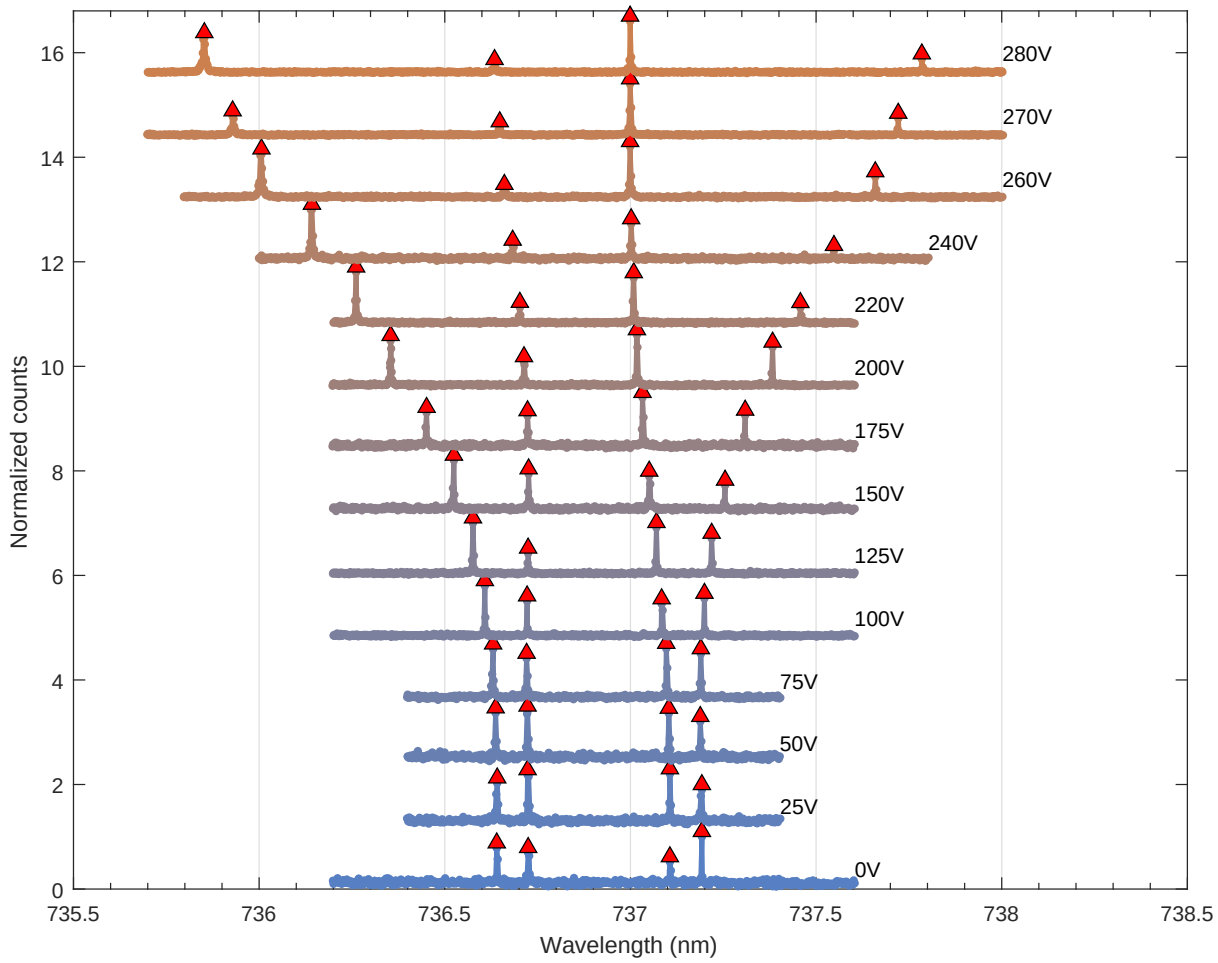

Figure S4: Tuning of PLE spectrum of a transverse-orientation SiV. Voltage applied to the device is indicated next to each spectrum.

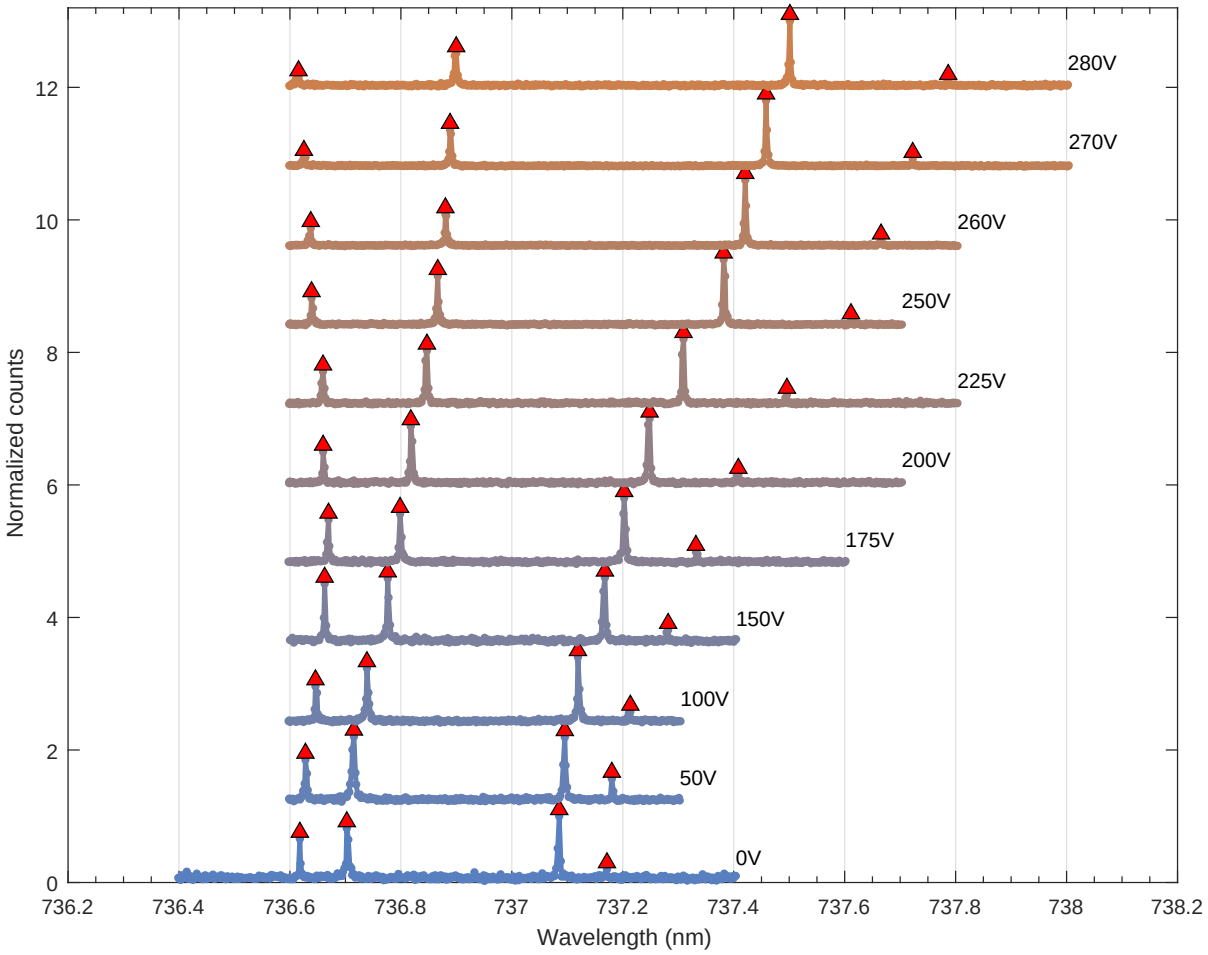

Figure S5: Tuning of PLE spectrum of an axial-orientation SiV. Voltage applied to the device is indicated next to each spectrum.

For both axial- and transverse-orientations, we rely on the inherent symmetries of the SiV to explain the strain-response of the electronic levels, and decompose the applied strain into A_{1g}- and E_g-modes, which are discussed in detail in Section 6.1. Briefly, we will show that transverse-orientation SiVs mainly experience E_g-strain, while axial-orientation SiVs experience both A_{1g}- and E_g-strain. E_g-strain separates A and D transitions apart by increasing the energy splitting between orbitals within the GS and ES as observed in Fig. S4. Meanwhile, A_{1g}-strain causes a linear shift in the mean wavelength of all four transitions. We extract the mean wavelength from the spectra of the axial-orientation SiV in Fig. S5, and plot it in Fig. S6 for greater clarity. In particular, with A_{1g}-strain, we show tuning of the mean wavelength by 0.3 nm (150 GHz), which is approximately 10 times the typical inhomogeneity in optical transition frequencies [8]. In the context of wavelength-tuning for cavity QED and photon-mediated entanglement of multiple emitters, the C line, which is the brightest and narrowest transition is typically of interest. We note that this transition is significantly more responsive to A_{1g}-strain (Fig. S5) than E_g-strain (Fig. S4).

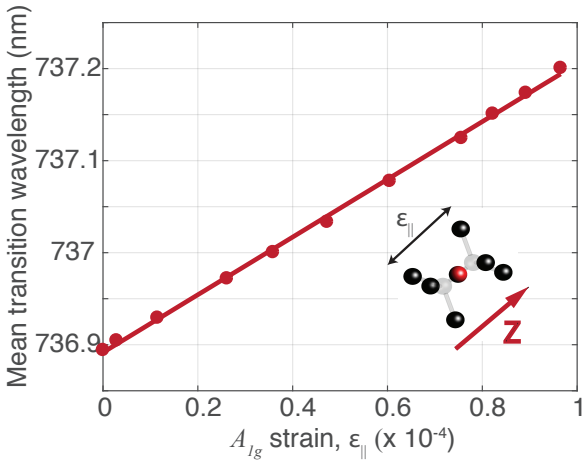

Figure S6: Tuning of mean optical wavelength of the axial-orientation SiV shown in Fig. S5. Solid line is a linear fit.

4.3 Stark effect

In this section, we evaluate the possibility of a Stark effect on SiV centres from DC electric field in our experimental setting. In our experiment, we apply voltages on the order of a few hundred volts across 4 μm gap, which induces an electric field on the order of tens of MV/m at the SiV location inside the cantilever. There is no first-order Stark effect for the SiV centre, since its inversion symmetry rules out a permanent dipole moment for the electronic levels [15]. A second-order Stark effect arises from the dipole moment between ground and excited states. The electric field response can be descibred by the following Hamiltonian ($h = 1$) written for any ground state $|g\rangle$ ($= |1\rangle, |2\rangle$), and any excited state $|e\rangle$ ($= |3\rangle, |4\rangle$).

$$\mathbb{H}^{\text{Stark}} = \begin{bmatrix} -\Delta_{\text{ZPL}}/2 & \mu_{eg} E_{\text{ext}} \\ \mu_{eg} E_{\text{ext}} & +\Delta_{\text{ZPL}}/2 \end{bmatrix} \tag{2}$$

where Δ_{ZPL} is the frequency of the relevant optical transitions (A,B,C or D), when there is no external perturbation. μ_{eg} is the dipole moment of the corresponding transition and E_{ext} is the applied electric field. To obtain an upper bound on the magnitude of the Stark shift, we assume that the external electric field is parallel to the dipole moment. The transition frequency as a function of the external electric field E_{ext} is given by $\Delta_{\text{ZPL}}(E_{\text{ext}}) \approx \Delta_{\text{ZPL}}(0) + 2\mu_{eg}^2 |E_{\text{ext}}|^2 / \Delta_{\text{ZPL}}(0)$.

A theoretical calculation of the transition dipole moment μ_{eg} of the SiV centre using density functional theory can be found in Ref. [16]. With the calculated value of $\mu_{eg}^2 \approx 42$ debye^2, we plot the expected second-order Stark shift for the C transition at $\Delta_{\text{ZPL}} = 406.7$ THz as a function of the external electric field in Fig. S7a. Next, we use FEM simulation to estimate the magnitude of electric field at the location of the SiV centre. Fig. S7b shows the colormap of electric field intensity in the cross section perpendicular to the cantilever's long axis, when a voltage of 200 V is applied. At the implantation depth of 50 nm, we estimate that SiVs will experience an electric field on the order of 10 MV/m at most, when a voltage on the order

of 100 V is applied to the device. Combining this information with the plot in Fig. S7a, the estimated Stark shift in our experiment is of the order of few hundred MHz, when a few hundred volts is applied. This is negligble, when compared with optical transition frequency shifts on the order of few hundred GHz observed for both axial and transverse SiVs in Fig. S6, and Fig. 2b of the main text respectively.

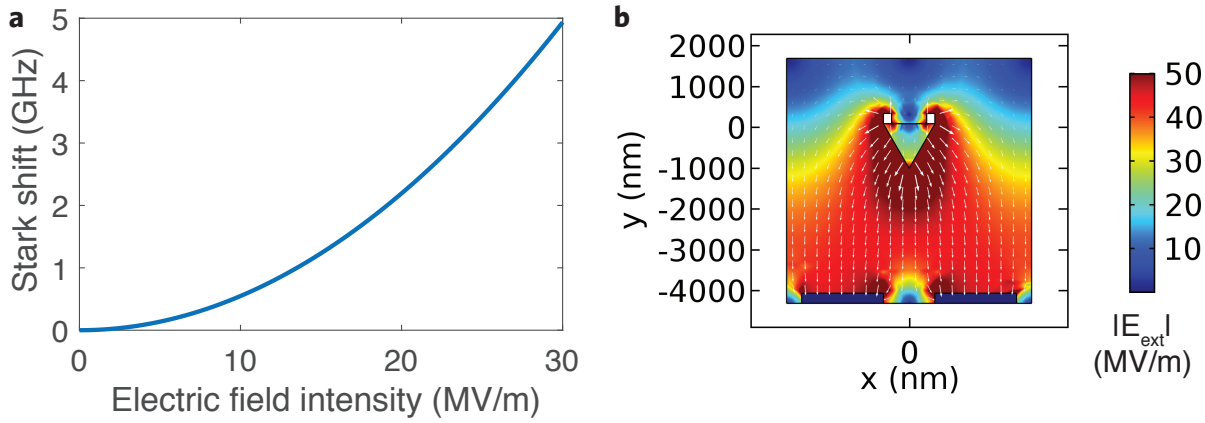

Figure S7: (a) Stark shift of the ZPL frequency as a function of applied electric field. (b) Cross-sectional plot of electric field intensity in the plan normal to the cantilever long axis for an applied voltage of 200 V. White arrows represent electric field vector in the plane with their length being proportional to the magnitude.

5 Strain-response theory and fitting

5.1 Group theory description of strain-response

The stress-response of the electronic levels of trigonal point-defects in cubic crystals was treated theoretically by Hughs and Runciman in [17]. A solution of this problem for the specific case of the SiV has been carried out using group theory in [18] with some errors. Here, we reconcile these two treatments, and present a model for the response of the SiV electronic levels to strain (and stress). In what follows, we use X, Y, Z to refer to the internal basis of the SiV (see inset of Fig. S3(a). eg. for a [111] oriented SiV, we have $X : [\bar{1}\bar{1}2], Y : [\bar{1}10], Z : [111]$), and x, y, z to refer to the axes of the diamond crystal, i.e. $x : [100], y : [010], z : [001]$. We use σ and ϵ for the stress and strain tensors in the SiV basis, and $\bar{\sigma}$ and $\bar{\epsilon}$ to refer to them in the crystal basis.

When the applied stress is small, the perturbation Hamiltonian $\mathbb{H}^{\text{strain}}$ can written in terms of stress-tensor components as discussed in [17]

$$\mathbb{H}^{\text{strain}} = \sum_{ij} A_{ij} \bar{\sigma}_{ij} \tag{3}$$

Here A_{ij} are operators for each stress-tensor component, and act on the SiV electronic levels.

i, j are indices for x, y and z. Stress and strain are linearly related via the elasticity tensor of the material c_{ijkl}.

$$\bar{\sigma}_{ij} = \sum_{kl} c_{ijkl} \bar{\epsilon}_{kl} \tag{4}$$

Thus the Hamiltonian 3 can also be written in terms of strain-tensor components as

$$\mathbb{H}^{\text{strain}} = \sum_{kl} B_{kl} \bar{\epsilon}_{kl} \tag{5}$$

where the operators B_{kl} are related to the operators A_{ij} according to

$$B_{kl} = \sum_{ij} c_{ijkl} A_{ij} \tag{6}$$

Group theory can be used to rewrite the strain Hamiltonian $\mathbb{H}^{\text{strain}}$ in terms of linear combinations of stress-tensor (strain-tensor) components adapted to the symmetries of the SiV center shown in the inset of Fig. S3(a). Each of these combinations can be viewed as a particular 'mode' of stress (strain). For instance, we will show subsequently in this section that $\bar{\sigma}_{xx} + \bar{\sigma}_{yy} + \bar{\sigma}_{zz}$ is one such mode - it is a fully symmetric deformation of the molecule. The effect of each mode on the electronic levels of the defect can be quite readily deduced using group theory, since the orbital wavefunctions corresponding to these electronic levels have specific symmetries themselves.

More technically, each stress (strain) mode transforms as an irreducible representation of the D_{3d} group, which is the point group of the SiV center. Specifically,

$$\mathbb{H}^{\text{strain}} = \sum_{r} A_r \bar{\sigma}_r \tag{7}$$

where r runs over all the modes. Each $\bar{\sigma}_r$ is a linear combination of the components $\bar{\sigma}_{ij}$, and corresponds to a specific irreducible representation. The electronic wavefunctions themselves correspond to particular irreducible representations (see section 2.1.4 of [18] for a derivation of the SiV orbitals, and their connection to irreducible representations). Therefore, deducing the matrix elements in each operator A_r simply requires computing the direct products of irreducible representations (see table 2.8 of [18] for direct products of irreducible representations of the D_{3d} group). Note that this group theory approach only allows one to predict the possibility of a non-zero matrix element in A_r, but does not give the values of these elements.

In section 2.2.5 of [18], it is established that the stress tensor transforms as the irreducible representation $A_{1g} + E_g$, which has even parity about the inversion center of the SiV. Since the ground states of the SiV transform as E_g (even), and the excited states transform as E_u (odd), stress does not couple the ground and excited states with each other to first order. As a result, we can describe the stress-response of the ground and excited state manifolds independently. In particular, $\mathbb{H}^{\text{strain}}$ is identical in form for both manifolds, but will involve different numerical values of strain-response coefficients. Therefore, we drop the subscripts g and u used to refer to the ground and excited states, and simply write the

interaction Hamiltonian in the doubly-degenerate basis of E-states $\{|e_X\rangle, |e_Y\rangle\}$. In [17], the strain-Hamiltonian $\mathbb{H}^{\text{strain}}$ is described using a decomposition of stress-tensor elements in the crystal-basis.

$$\mathbb{H}^{\text{strain}} = \begin{bmatrix} \alpha - \beta & \gamma \\ \gamma & \alpha + \beta \end{bmatrix} \tag{8}$$

The components α, β, γ are given by the linear combinations

$$\begin{aligned} \alpha &= A_1(\bar{\sigma}_{xx} + \bar{\sigma}_{yy} + \bar{\sigma}_{zz}) + 2A_2(\bar{\sigma}_{yz} + \bar{\sigma}_{zx} + \bar{\sigma}_{xy}) \\ \beta &= B(2\bar{\sigma}_{zz} - \bar{\sigma}_{xx} - \bar{\sigma}_{yy}) + C(2\bar{\sigma}_{xy} - \bar{\sigma}_{yz} - \bar{\sigma}_{zx}) \\ \gamma &= \sqrt{3}B(\bar{\sigma}_{xx} - \bar{\sigma}_{yy}) + \sqrt{3}C(\bar{\sigma}_{yz} - \bar{\sigma}_{zx}) \end{aligned} \tag{9}$$

The four Hughes-Runciman coefficients A_1, A_2, B, C completely determine the stress-response of the $\{|e_X\rangle, |e_Y\rangle\}$ manifold. The two terms multiplied by A_1 and A_2 that make up α are stress-modes that transform as A_{1g}, and lead to a common mode energy shift of the $\{|e_X\rangle, |e_Y\rangle\}$ orbitals. On the other hand, the terms in β and γ multiplied by B and C are stress-modes that transform as E_g. They change the energy-splitting, and mix the orbitals respectively.

To gain more physical intuition for these stress-modes, we can write α, β, γ in terms of the stress-tensor in the SiV basis σ. σ and $\bar{\sigma}$ are related by the unitary transformation $R = R_z(45°)R_y(54.7°)$, where $R_z(\theta)$, and $R_x(\phi)$ correspond to rotations by θ and ϕ about the z- and x-axes respectively.

$$\bar{\sigma} = R\sigma R^T$$

Substituting expressions for individual components of $\bar{\sigma}$ from the above formula in equations (9), we get

$$\begin{aligned} \alpha &= (A_1 - A_2)(\sigma_{XX} + \sigma_{YY}) + (A_1 + 2A_2)\sigma_{ZZ} \\ \beta &= (B + C)(\sigma_{XX} - \sigma_{YY}) + \sqrt{2}(C - 2B)\sigma_{ZX} \\ \gamma &= -2(B + C)\sigma_{XY} + \sqrt{2}(C - 2B)\sigma_{YZ} \end{aligned} \tag{10}$$

Table 1 describes the effects of various modes of stress on the $\{|e_X\rangle, |e_Y\rangle\}$ orbitals, and the stress-response coefficients (susceptibilities) of these modes.

Stress term	Effect	Susceptibility in terms of Hughs-Runciman coefficients
$\sigma_{XX} + \sigma_{YY}$	Common-mode shift	$(A_1 - A_2)$
σ_{ZZ}	Common-mode shift	$(A_1 + 2A_2)$
$\sigma_{XX} - \sigma_{YY}$	Relative shift	$(B + C)$
σ_{ZX}	Relative shift	$\sqrt{2}(C - 2B)$
σ_{XY}	Mixing	$-2(B + C)$
σ_{YZ}	Mixing	$\sqrt{2}(C - 2B)$

Table 1: Various stress-modes, and their effects on the $\{|e_X\rangle, |e_Y\rangle\}$ orbitals.

We now use equation (4) to write the relations (9) in terms of strain-tensor components. For notational brevity, we can condense pairs of indices ij as $xx \to 1, yy \to 2, zz \to 3, xy \to 4, \ldots$. Crystals with cubic symmetry only have non-zero c_{11}, c_{22}, c_{44} [19]. This gives us the following stress-strain relations.

$$\begin{aligned} \bar{\sigma}_{ii} &= c_{11}\bar{\epsilon}_{ii} + \sum_{j\neq i} c_{12}\bar{\epsilon}_{jj} \\ \bar{\sigma}_{ij} &= c_{44}\bar{\epsilon}_{ij} \end{aligned} \tag{11}$$

For diamond, we have $c_{11} = 1075\,\text{GPa}$, $c_{12} = 139\,\text{GPa}$, $c_{44} = 567\,\text{GPa}$ [19]. Substituting the relations (11) in equations (9), we obtain the following strain-relations similar in form to the stress-relations, but with different coefficients.

$$\begin{aligned} \alpha &= \mathscr{A}_1(\bar{\epsilon}_{xx} + \bar{\epsilon}_{yy} + \bar{\epsilon}_{zz}) + 2\mathscr{A}_2(\bar{\epsilon}_{yz} + \bar{\epsilon}_{zx} + \bar{\epsilon}_{xy}) \\ \beta &= \mathscr{B}(2\bar{\epsilon}_{zz} - \bar{\epsilon}_{xx} - \bar{\epsilon}_{yy}) + \mathscr{C}(2\bar{\epsilon}_{xy} - \bar{\epsilon}_{yz} - \bar{\epsilon}_{zx}) \\ \gamma &= \sqrt{3}\mathscr{B}(\bar{\epsilon}_{xx} - \bar{\epsilon}_{yy}) + \sqrt{3}\mathscr{C}(\bar{\epsilon}_{yz} - \bar{\epsilon}_{zx}) \end{aligned} \tag{12}$$

The coefficients $\mathscr{A}_1, \mathscr{A}_2, \mathscr{B}, \mathscr{C}$ in terms of the Hughes-Runciman coefficients A_1, A_2, B, C are given by

$$\begin{aligned} \mathscr{A}_1 &= c_{11}A_1 \\ \mathscr{A}_2 &= c_{12}A_1 + c_{44}A_2 \\ \mathscr{B} &= c_{11}B \\ \mathscr{C} &= c_{44}C - c_{12}B \end{aligned} \tag{13}$$

Just as we did for the stress-relations, we can rotate the strain tensor $\bar{\epsilon}$ into the SiV basis to get relations similar to those in (10). We write these below.

$$\begin{aligned} \alpha &= t_\perp(\epsilon_{XX} + \epsilon_{YY}) + t_\parallel \epsilon_{ZZ} \\ \beta &= d(\epsilon_{XX} - \epsilon_{YY}) + f\epsilon_{ZX} \\ \gamma &= -2d\epsilon_{XY} + f\epsilon_{YZ} \end{aligned} \tag{14}$$

Here $t_\perp, t_\parallel, d, f$ are the four strain-susceptibility parameters that completely describe the strain-response of the $\{|e_X\rangle, |e_Y\rangle\}$ manifold. Table 2 shows the strain-susceptibilities in terms of coefficients used earlier in this section, and defines the effect of corresponding strain-modes on the $\{|e_X\rangle, |e_Y\rangle\}$ manifold.

At this juncture, we contrast equations (14) with the results in [18] (equations 2.80-2.82). Our analysis, which follows from the original treatment by Hughes and Runciman in [17] captures the strain-modes derived in [18], but in addition, predicts a non-zero response to axial strain ϵ_{ZZ}, and the shear strains ϵ_{ZX} and ϵ_{YZ}.

Strain term	Effect	Susceptibility	Relation to Hughes-Runciman coefficients
$\epsilon_{XX}+\epsilon_{YY}$	Common-mode shift	$t_\perp$	$\mathscr{A}_1-\mathscr{A}_2=c_{11}A_1-(c_{12}A_1+c_{44}A_2)$
ϵ_{ZZ}	Common-mode shift	$t_\parallel$	$\mathscr{A}_1+2\mathscr{A}_2=c_{11}A_1+2(c_{12}A_1+c_{44}A_2)$
$\epsilon_{XX}-\epsilon_{YY}$	Relative shift	d	$\mathscr{B}+\mathscr{C}=(c_{11}-c_{12})B+c_{44}C$
ϵ_{XY}	Mixing	$-2d$	
ϵ_{ZX}	Relative shift	f	$\mathscr{C}-2\mathscr{B}=c_{44}C-(c_{12}+2c_{11})B$
ϵ_{YZ}	Mixing	f	

Table 2: Various strain-modes, and their effects on the $\{|e_X\rangle, |e_Y\rangle\}$ orbitals.

5.2 Estimation of strain-susceptibility parameters

Without any external perturbation such as strain or external magnetic field, spin-orbit coupling gives the following interaction Hamiltonian within ground and excited state manifolds, expressed in the $\{|e_X\rangle, |e_Y\rangle\} \otimes \{|\uparrow\rangle, |\downarrow\rangle\}$ basis [18].

$$\mathbb{H}^{\text{SO}} = -\frac{\lambda_{SO}}{2}\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \otimes \sigma_z \tag{15}$$

Here, λ_{SO} is the spin-orbit coupling strength within each manifold (46 GHz for the ground states, and 255 GHz for the excited states). σ_z is the $S = 1/2$ Pauli matrix that takes the spin degree of freedom into account. The strain Hamiltonian (8) after including the spin degree of freedom can be expressed as

$$\mathbb{H}^{\text{strain}} = \begin{bmatrix} \alpha-\beta & \gamma \\ \gamma & \alpha+\beta \end{bmatrix} \otimes \mathbb{I}_2 \tag{16}$$

Note that identity matrix for the spin in equation (16) implies that strain affects purely the orbital part of the wavefunction. Combining equations (15) and (16), we get the following total Hamiltonian.

$$\mathbb{H}^{\text{total}} = \mathbb{H}^{\text{SO}} + \mathbb{H}^{\text{strain}} = \begin{bmatrix} \alpha-\beta & 0 & \gamma-i\lambda_{SO}/2 & 0 \\ 0 & \alpha-\beta & 0 & \gamma+i\lambda_{SO}/2 \\ \gamma+i\lambda_{SO}/2 & 0 & \alpha+\beta & 0 \\ 0 & \gamma-i\lambda_{SO}/2 & 0 & \alpha+\beta \end{bmatrix} \tag{17}$$

Diagonalization of the total Hamiltonian gives the following two distinct eigenvalues within ground and excited state manifolds, respectively. Each the energy eigenstate is doubly degenerate due to the spin degree of freedom.

$$E_1 = \frac{1}{2}\left(2\alpha - \sqrt{\lambda_{SO}^2+4\beta^2+4\gamma^2}\right), \quad E_2 = \frac{1}{2}\left(2\alpha + \sqrt{\lambda_{SO}^2+4\beta^2+4\gamma^2}\right) \tag{18}$$

Equation (18) is valid within both ground and excited state manifolds. Therefore, we can extract following quantities that can be directly measured via spectroscopy.

$$\Delta_{\text{ZPL}} = \Delta_{\text{ZPL},0} + (\alpha_e - \alpha_g)$$

$$= \Delta_{\mathrm{ZPL},0} + \left(t_{\parallel,e} - t_{\parallel,g}\right)\epsilon_{ZZ} + \left(t_{\perp,e} - t_{\perp,g}\right)\left(\epsilon_{XX} + \epsilon_{YY}\right) \tag{19}$$

$$\Delta_{\mathrm{gs}} = \sqrt{\lambda_{SO,g}^2 + 4\left(\beta_g^2 + \gamma_g^2\right)}$$

$$= \sqrt{\lambda_{SO,g}^2 + 4\left[d_g(\epsilon_{XX} - \epsilon_{YY}) + f_g\epsilon_{YZ}\right]^2 + 4\left[-2d_g\epsilon_{XY} + f_g\epsilon_{ZX}\right]^2} \tag{20}$$

$$\Delta_{\mathrm{es}} = \sqrt{\lambda_{SO,e}^2 + 4\left(\beta_e^2 + \gamma_e^2\right)}$$

$$= \sqrt{\lambda_{SO,e}^2 + 4\left[d_e(\epsilon_{XX} - \epsilon_{YY}) + f_e\epsilon_{YZ}\right]^2 + 4\left[-2d_e\epsilon_{XY} + f_e\epsilon_{ZX}\right]^2} \tag{21}$$

Here, the subscript $g(e)$ refers to the ground (excited) state manifold. Δ_{ZPL} is the mean zero phonon line (ZPL) frequency, and Δ_{gs} (Δ_{es}) is the ground (excited) state splitting, respectively. $\Delta_{\mathrm{ZPL},0}$ is the mean ZPL frequency, when there is no external strain. From these equations, we directly observe that strain of A_{1g} type, which determines the α terms, leads to a change in the mean ZPL frequency. On the other hand, strain of E_g type, which determines β, γ terms, leads to a change in the GS and ES splittings.

All three values, Δ_{ZPL}, Δ_{gs}, and Δ_{es} as a function of strain can be directly extracted from the photoluminescence excitation (PLE) spectra shown in Figs. S4 and S5. Fitting the model equations (19-21) to the experimental data, we should be able to estimate the strain-susceptibility parameters. To extract all the values $\{t_\perp, t_\parallel, d, f\}$ for both ground and excited state manifolds, in principle, strain needs to be applied at least in three different directions for a given SiV. This procedure gives a set of overdetermined equations in these parameters [17]. However, the devices in this study can only induce two types of strain profiles as shown in Fig. S3(b) and (c). In particular, for a given SiV in either the 'axial' or the 'transverse' class, the relative ratio between strain-tensor components remains constant, when the voltage applied to the device is swept. This condition makes it difficult to estimate the individual contributions of the $t_\parallel$ and $t_\perp$ terms to α, and of the d and f terms to β and γ.

To get around this issue, we follow an approximate approach. From Fig. S3(b), we observe that in the case of an axial SiV, $\epsilon_{ZZ} \gg (\epsilon_{XX} + \epsilon_{YY})$ is always true. Therefore, we can use the response of the axial SiV in Fig. S5 to approximately estimate $\left(t_{\parallel,e} - t_{\parallel,g}\right)$ by neglecting $(\epsilon_{XX} + \epsilon_{YY})$ in equation (19). Fig. S6 plots the mean ZPL frequency of the axial SiV in Fig. S5 vs. strain estimated from FEM simulation as described in Section 3. The slope of the linear fit yields $\left(t_{\parallel,e} - t_{\parallel,g}\right)$.

$$\left(t_{\parallel,e} - t_{\parallel,g}\right) = 1.7\,\mathrm{PHz/strain} \tag{22}$$

Likewise, in the case of the transverse SiV in Fig. S3(c), we can conclude that $(\epsilon_{XX} - \epsilon_{YY}) \gg \max\{\epsilon_{ZX}, \epsilon_{YZ}\}$. With this class of SiVs, we can approximately estimate $\{d_g, d_e\}$ by neglecting $\{\epsilon_{ZX}, \epsilon_{YZ}\}$ in equations (20,21). Fig. 2(b) of the main text plots the GS and ES splittings of the transverse SiV in Fig. S4 vs. strain estimated from FEM simulation as described in Section 3. Fitting yields

$$d_g = 1.3, \qquad d_e = 1.8 \qquad \mathrm{PHz/strain} \tag{23}$$

Once we extract $\left(t_{\parallel,e} - t_{\parallel,g}\right)$ from an axial SiV, we can use this value to further extract

$(t_{\perp,e} - t_{\perp,g})$ by fitting equation (19) to the tuning behavior of the mean ZPL frequency of the transverse SiV. This procedure yields

$$(t_{\perp,e} - t_{\perp,g}) = 78\,\text{THz/strain} \tag{24}$$

We immediately note that $(t_{\parallel,e} - t_{\parallel,g})$ is more than an order of magnitude larger than $(t_{\perp,e} - t_{\perp,g})$. This implies that ϵ_{ZZ} tunes the ZPL much more effectively than $(\epsilon_{XX} + \epsilon_{YY})$. This can be intuitively explained by examining the spatial profile of the GS and ES orbitals (Table 2.7 of [18]). Since the GS and ES are even (g) and odd (u) eigenstates of SiV's D_{3d} point symmetry group respectively, the charge density distributions of the orbitals e_{gX}, e_{uX} (and e_{gY}, e_{uY}) are similar in XY plane of nonzero Z, as it can be confirmed from the orbital pictures in [18]. As a result, we would expect that the common mode energy shift resulting from the strain-mode $\epsilon_{XX} + \epsilon_{YY}$ is very similar for the GS and ES manifolds, i.e. $t_{\perp,g} \approx t_{\perp,e}$. On the other hand, the energy shift from ϵ_{ZZ} is expected to have the opposite signs for the GS and ES manifolds due to the change in wavefunction parity along the Z-axis.

As the last step, we extract the values f_g, f_e in equations (20,21). We observe from table 2 that knowledge of d and B can allow us to determine f. We note that in section 5.1.1 of [18], the stress-response coefficients B_g=484 GHz/GPa and B_e=630 GHz/GPa are extracted based on uniaxial stress measurements carried out in [20]. Combining our estimates of d_g and d_e with this information, we predict

$$f_g = -250, \qquad f_e = -720 \qquad \text{THz/strain} \tag{25}$$

We emphasize that the strain-susceptibility values estimated in this section are subject to errors arising from (i) imprecision in SiV depth from the diamond surface (10% straggle from SRIM calculations, and in practice, higher due to ion-channeling effects), and (ii) due to the fact that the device geometry cannot be replicated exactly in the FEM simulation for strain estimation. In particular, the values f and $t_\perp$ are subject to large errors, since the E_g and A_{1g} responses are mostly dominated by the susceptibilities d and $t_\parallel$ respectively.

6 High-strain PLE spectra

In our devices, we can generate uniaxial strain as high as $\sim 10^{-3}$ at a voltage of 600 V. Example spectra showing strain-tuning into this regime are shown in Fig. S8. At high strain, as $\Delta_{\text{gs}} \gg k_B T \sim$80 GHz at T=4 K, population in the upper branch $|2\rangle$ of the ground-state manifold decreases exponentially, while the population in the lower branch $|1\rangle$ correspondingly increases to near-unity. As a result, with increasing strain, the B and D transitions become weaker in intensity, and eventually vanish. Simultaneously, the linewidth of the A transition increases owing to increasingly rapid phonon-mediated relaxation in the excited state from $|4\rangle$ to $|3\rangle$. This is captured by the model for single-phonon emission and absorption discussed in Section 9. On the other hand, as the C transition connects the lower orbitals in ground and excited states ($|1\rangle$ and $|3\rangle$), its linewidth is not affected as strongly by phonon-mediated processes, and is measured to be relatively unchanged with strain (See Section 12). The C transition also becomes brighter at high-strain due to near-unity population in the lower

ground state branch $|1\rangle$.

While the high-strain spectra only reveal A and C transitions, the difference in their frequencies gives us the excited-state splitting Δ_{es} exactly. Since the same strain-components are responsible for increasing Δ_{gs} and Δ_{es}, we can use the strain-susceptibilities extracted in Section 6.2 to estimate Δ_{gs}. Using this procedure, for the highest strain condition in Fig. S8, we infer $\Delta_{gs} = 1.2\,\text{THz}$. A more precise experimental technique to measure the GS- and ES-splittings at high strain is described in Section 10.4.

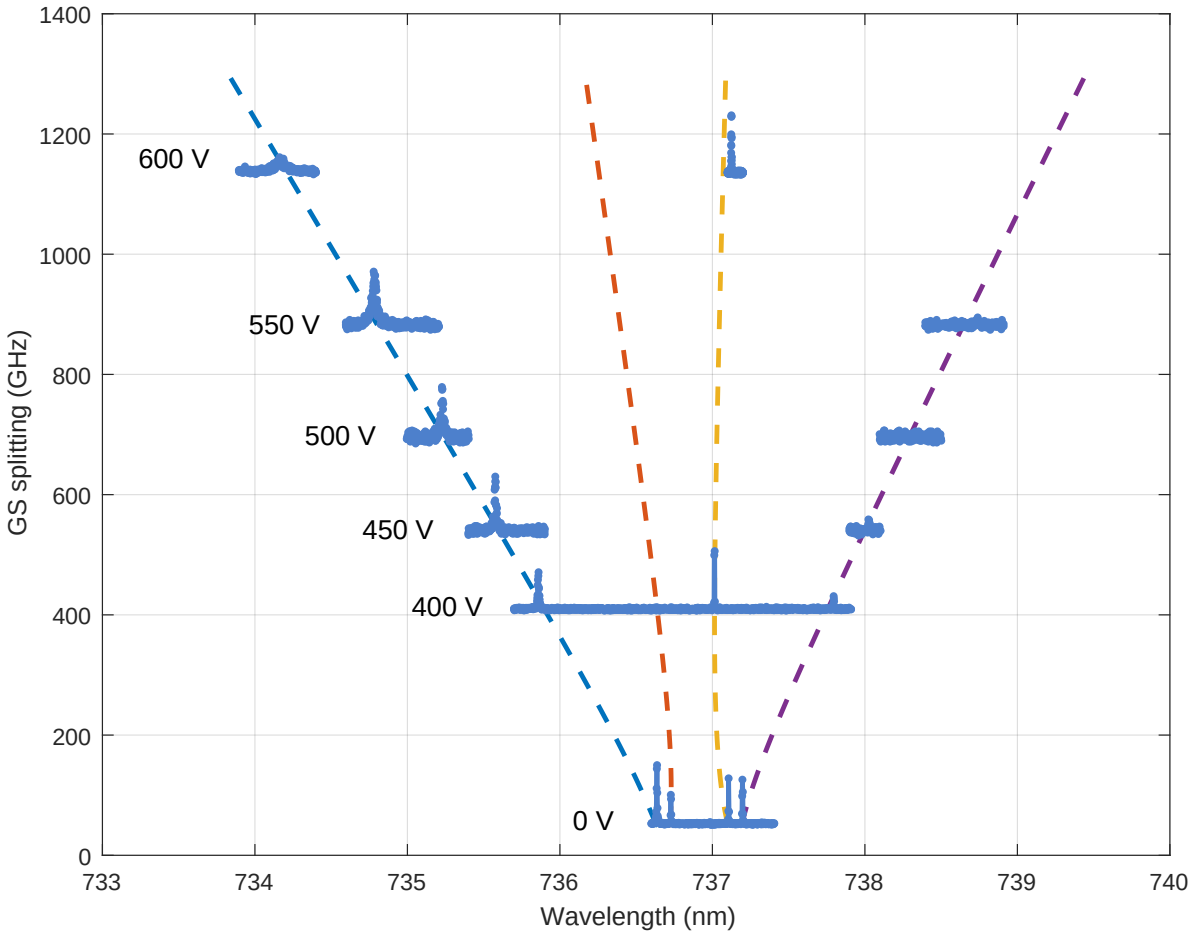

Figure S8: Tuning of PLE spectrum of a transverse SiV subject to high strain. Voltage applied to the device is indicated next to each spectrum. y-axis shows ground-state splitting Δ_{gs} corresponding to the spectrum estimated using the procedure detailed in the text. Dashed lines correspond to modeled strain-response of the four optical transitions.

7 Orbital thermalization measurements

We use time resolved pump-probe fluorescence to characterize the phonon processes in the GS. In this method, two consecutive laser pulses resonant with the D transition are used to, first initialise GS orbital population in the lower branch $|1\rangle$, and after a set delay τ, read-out population in the upper branch $|2\rangle$. A schematic of the pulse sequence, and an example of a resulting fluorescence time-trace are shown in Fig. S9. By repeating this sequence for steadily increasing pump-probe delay τ, we measure the rate at which the GS population relaxes towards thermal equilibrium due to resonant phonons.

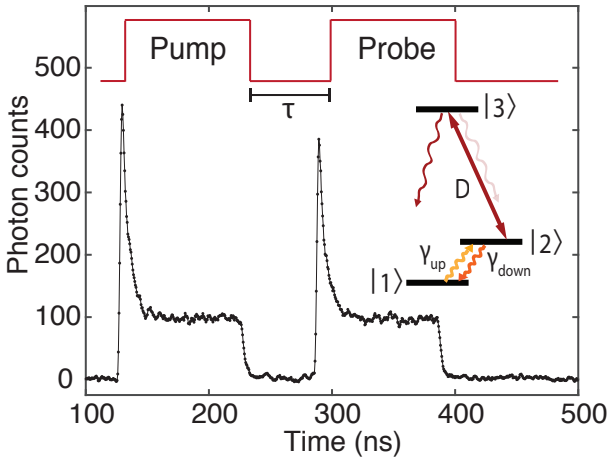

Figure S9: Time-resolved fluorescence signal in pump-probe measurement for a delay $\tau = 50$ ns between the two laser pulses. The laser is resonant with the D transition, and optically pumps the GS population into the lower orbital branch $|1\rangle$ over a timescale of few ns. After time τ, the fluorescence signal from the probe pulse has a leading edge determined by the population in the upper orbital branch $|2\rangle$. The decay rates between levels $|1\rangle$ and $|2\rangle$ - γ_{up} due to phonon-absorption, and γ_{down} due to phonon-emission - are also shown.

7.1 Experimental setup

The pump-probe pulse sequence described above is implemented by pulsing our resonant-excitation laser with a Mach-Zehnder intensity electro-optic modulator (EO Space AZ-AV5-5-PFA-PFA-737) driven by a digital-delay generator with rise- and fall-times of 2 ns (SRS DG645). Over the course of the measurements, the operation point of the intensity electro-optic modulator (EOM) is stabilized against long-term drifts with continuous feedback on the DC-bias voltage. The feedback loop is implemented with a lock-in amplifier (SRS SR830) generating a low-frequency (1 KHz) modulation of the DC-bias voltage. Photon-count pulses from the single-photon-detector are time-tagged on a PicoHarp 300 module triggered by the delay-generator. The laser frequency itself is stabilized by continuous feedback with a wavemeter (High Finesse WS7).

7.2 Extraction of thermalization rate

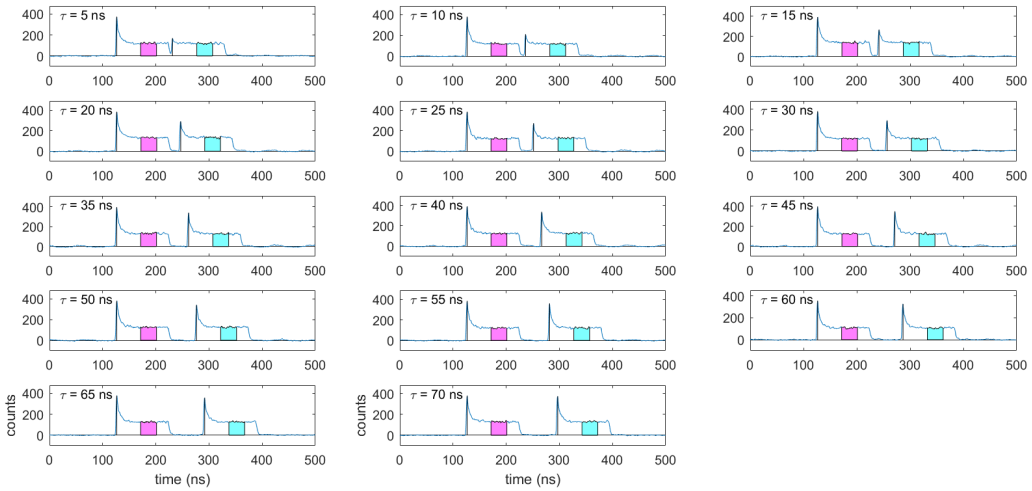

Figure S10: Fluorescence time-traces for various pump-probe delays between τ=5 ns to 70 ns taken at GS-splitting $\Delta_{\rm gs}$=46 GHz. x-axis is time in ns, and y-axis is photon counts integrated over multiple iterations of the pulse sequence.

Example data from implementing the pulse sequence in Fig. S9 for various pump-probe delays is shown in Fig. S10. This data can be interpreted and processed to yield a GS-population thermalization curve as follows. As shown in Fig. S11a, the leading edge of the first fluorescence signal corresponds to thermal population $p_{2,th}$ in the GS level $|2\rangle$. Upon switching on the pump pulse, this decays to a residual value $p_{2,opt}$ determined by the competition between the optical pumping rate (above saturation, this is simply the decay rate γ_e from the excited state $|3\rangle$) and the rates $\gamma_{\rm up}, \gamma_{\rm down}$. After time delay τ, the leading edge of the probe fluorescence signal corresponds to partially recovered population $p_2(\tau)$ due to thermalization. We can describe the population recovery in level $|2\rangle$ as

$$(p_2(\tau) - p_{2,th}) = (p_2(\tau) - p_{2,opt}) \exp\left[-\left(\gamma_{\rm up} + \gamma_{\rm down}\right) t\right] \tag{26}$$

In particular, we calculate the normalized change in photon-counts, $(p_2(\tau) - p_{2,th}) \,/\, (p_2(\tau) - p_{2,opt})$ from each measurement in Fig. S10, and carry out an exponential fit in Fig. S11b to extract $(\gamma_{\rm up} + \gamma_{\rm down})$. Repeating this experiment for various values of GS-splitting $\Delta_{\rm gs}$, we arrive at Fig. 2c in the main text.

From the optical-pumping fluorescence signal, we can define a pumping efficiency $\left(1 - \frac{p_{2,opt}}{p_{2,th}}\right)$. This describes the ability of the pump pulse to initialize GS population in level $|1\rangle$ by acting against the thermalization process. As the GS-splitting $\Delta_{\rm gs}$ is increased, we find that the pumping-efficiency decreases in Fig. S12. This can be attributed to the downward thermalization rate $\gamma_{\rm down}$ rapidly increasing with $\Delta_{\rm gs}$ as will be discussed in Section 9. At large $\Delta_{\rm gs}$, the optical pumping rate, R_{opt}, which is $\sim \gamma_e$, the excited state decay rate cannot substantially outweigh $\gamma_{\rm down}$. As a result, the GS population cannot be optically pumped

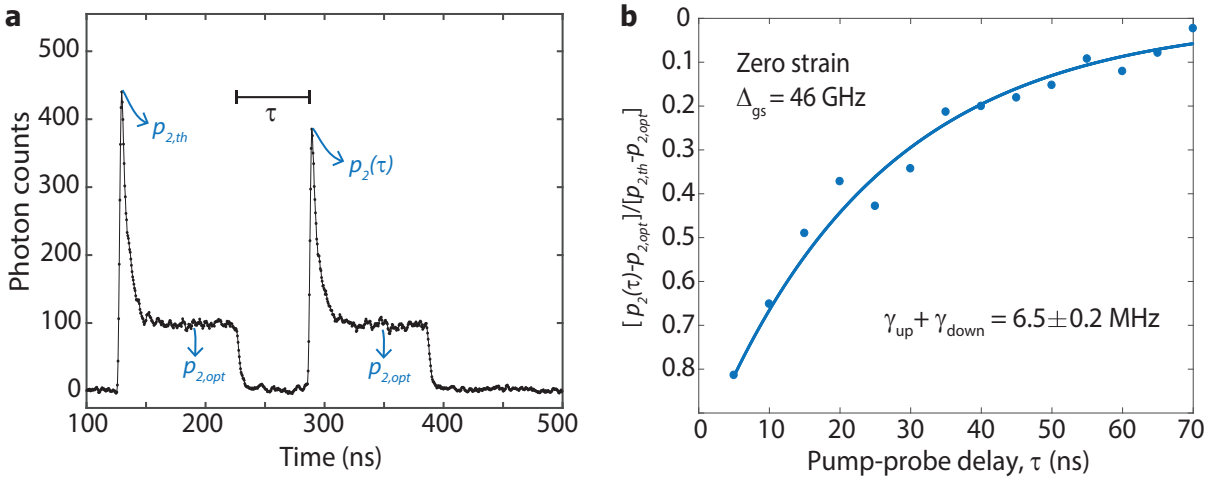

Figure S11: (a) Fluorescence time-trace for τ=50 ns from Fig. S10 showing relevant quantities related to the population in level $|2\rangle$. (b) Thermalization curve constructed by extracting the normalized change in photon-counts for various pump-probe delays τ. Solid line is an exponential fit.

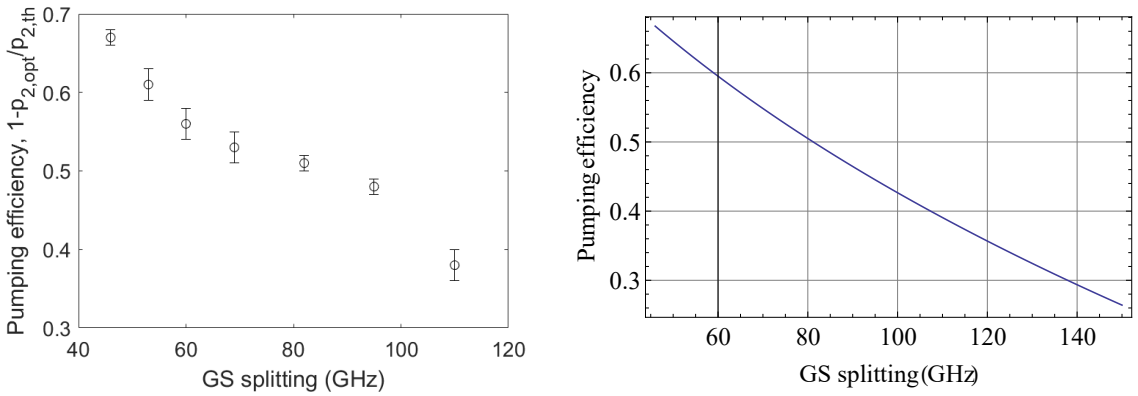

Figure S12: (a) Experimentally measured optical-pumping efficiency (see text for definition) versus GS-splitting Δ_{gs}. (b) Theoretically estimated optical-pumping efficiency using a three-level rate equations model. We assume optical pumping above saturation, GS thermalization rates, γ_{up} and γ_{down} to follow the model in Section 9, excited state decay rate γ_e = 90 MHz, branching ratio into $|1\rangle$ = 0.5.

into level $|1\rangle$ better than is dictated by thermal equilibrium. This makes it impractical to measure the thermalization rate beyond a certain Δ_{gs} using the pump-probe technique. In our experiments, we measure up to Δ_{gs} = 110 GHz.

8 Orbital thermalization model

8.1 Theory

Due to the nature of the strain interaction (Eq. 8), phonons having non-zero E_g-strain components that are resonant with the ground state splitting Δ_{gs} can directly drive the orbital transition. Indeed, in [21], single-phonon processes are found to be the dominant cause for thermalization of the ground state orbitals at 4 K. We follow the approach used in this work to model the orbital thermalization rate due to single-phonon processes. Summing over all acoustic modes of polarizations λ, and wave-vectors $\mathbf{k}$, we can write the upward and downward orbital relaxation rates as

$$\gamma_{\text{up}}(\Delta_{\text{gs}}) = 2\pi \sum_{\lambda,\mathbf{k}} n_{th}(\nu)|g(\lambda,\mathbf{k})|^2 \delta(\nu - \Delta_{\text{gs}}) \tag{27}$$

$$\gamma_{\text{down}}(\Delta_{\text{gs}}) = 2\pi \sum_{\lambda,\mathbf{k}} (n_{th}(\nu)+1)|g(\lambda,\mathbf{k})|^2 \delta(\nu - \Delta_{\text{gs}}) \tag{28}$$

Here $g(\lambda,\mathbf{k})$ is the single-phonon coupling rate to the mode with polarization λ and wave-vector $\mathbf{k}$, and $n_{th}(\nu)$ is the Bose-distribution.

$$n_{th}(\nu) = \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \tag{29}$$

Note that the term $+1$ in the downward rate in equation (28) physically corresponds to spontaneous emission of a phonon, a process that is independent of the thermal occupation of the modes.

Consider a longitudinal mode with uniaxial zero-point strain ϵ_{ZPF} perfectly aligned to drive the orbital transition with susceptibility d. In this case, we would have $g(\lambda,\mathbf{k}) = 2\pi d\epsilon_{ZPF}$. ϵ_{ZPF} itself can be derived by assuming an elastic energy of $h\nu$ localized in a mode volume V.

$$h\nu = \frac{1}{2} E \epsilon_{ZPF}^2 V$$

where E is the elastic modulus. This gives

$$g(\lambda,\mathbf{k}) = 2\pi d\sqrt{\frac{2h\nu}{EV}} \tag{30}$$

Regardless of the exact details of the mode profile, the relation $g \propto \sqrt{\nu}$ will hold true. In general, the proportionality constant will include a geometric factor, relevant components of the elasticity tensor, and a combination of the susceptibilities d and f, which describe mixing of the $|e_X\rangle, |e_Y\rangle$ orbitals. Neglecting these details and further assuming an isotropic dispersion relation, we define an average single-phonon coupling rate $\overline{g(\lambda,\mathbf{k})}$ with

$$|\overline{g(\lambda,\mathbf{k})}|^2 = \chi\nu \tag{31}$$

Converting equations (27) and (28) into integrals in 3D k-space, we get

$$\begin{aligned}\gamma_{\text{up}}(\Delta_{\text{gs}}) &= 2\pi \int |\overline{g(\lambda,\mathbf{k})}|^2 n_{th}(\nu)\delta(\nu-\Delta_{\text{gs}})4\pi k^2 dk \\ &= 2\pi \int \chi\nu \cdot n_{th}(\nu)\delta(\nu-\Delta_{\text{gs}})\rho\nu^2 d\nu\end{aligned}$$

Here $\rho\nu^2$ gives the density of states (DOS) for phonons in the bulk, where ρ is a constant that depends on the speed of sound averaged over various modes. Thus we get

$$\gamma_{\text{up}}(\Delta_{\text{gs}}) = 2\pi\chi\rho\Delta_{\text{gs}}^3 n_{th}(\Delta_{\text{gs}}) \tag{32}$$

Similarly, we can write

$$\gamma_{\text{down}}(\Delta_{\text{gs}}) = 2\pi\chi\rho\Delta_{\text{gs}}^3 (n_{th}(\Delta_{\text{gs}})+1) \tag{33}$$

In these expressions, the first term in the product, $2\pi\chi\rho\Delta_{\text{gs}}^3$ corresponds to the mean-squared single-phonon coupling rate multiplied with the DOS at the GS splitting Δ_{gs}, while the second term corresponds to the thermal occupation of each mode. Fig. S13 shows plots of the calculated dependence of upward and downward rates on Δ_{gs} at temperature $T = 4\,\text{K}$. We observe that the upward rate shows a non-monotonic behavior, approaching its maximum value around $h\Delta_{\text{gs}} \sim k_B T$. The increasing DOS term dominates in the regime $h\Delta_{\text{gs}} < k_B T$, and causes γ_{up} to increase. However, when $h\Delta_{\text{gs}} \gg k_B T$, the thermal occupation of the modes behaves as $n_{th}(\Delta_{\text{gs}}) = \exp\left(-\frac{h\Delta_{\text{gs}}}{k_B T}\right)$. This exponential roll-off dominates the polynomially increasing DOS, and causes γ_{up} to decrease at higher strain. In contrast, the downward rate monotonically increases with the GS-splitting, because it is dominated by the spontaneous emission rate, which simply scales as the DOS.

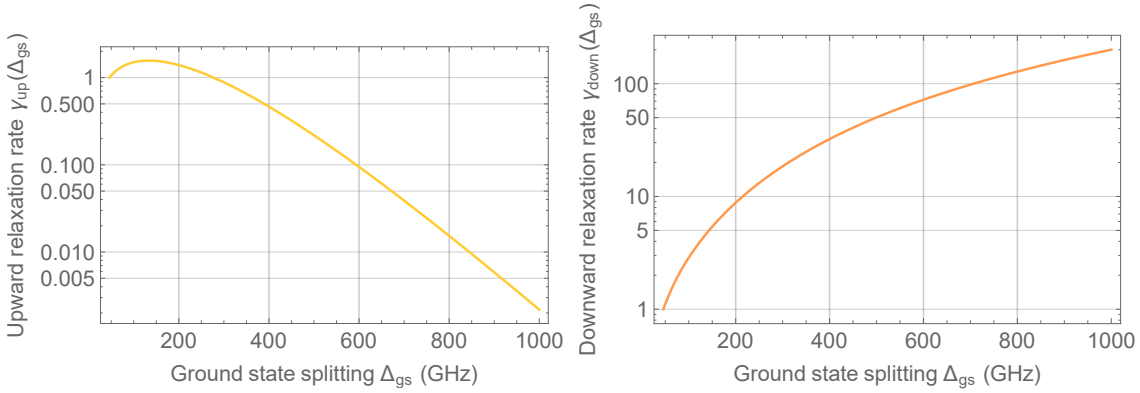

Figure S13: Variation of the upward (phonon-absorption) and downward (phonon-emission) relaxation rates with GS-splitting Δ_{gs} at temperature T=4 K. The y-axis is normalized to the zero-strain rates $\gamma_{\text{up}}(46\text{GHz})$ and $\gamma_{\text{down}}(46\text{GHz})$ in each plot respectively.

8.2 Fitting

With the pump-probe technique used in our experiments, we measure the sum of the upward and downward rates, which should have the dependence

$$\gamma_{\mathrm{up}}(\Delta_{\mathrm{gs}}) + \gamma_{\mathrm{down}}(\Delta_{\mathrm{gs}}) = A\Delta_{\mathrm{gs}}^3(2n_{th}(\Delta_{\mathrm{gs}}) + 1) = A\Delta_{\mathrm{gs}}^3 \coth\left(\frac{h\Delta_{\mathrm{gs}}}{k_B T}\right) \tag{34}$$

where A is a constant. Our cantilever has lateral dimensions of order $1\,\mu$m. This is about an order of magnitude larger than half the acoustic wavelength for $\Delta_{\mathrm{gs}} = 50\,$GHz in diamond, which is $\sim$120 nm. Thus, the cantilever itself is a bulk-like structure for the SiV. However, as can be observed in the left panel of Fig. S14 our experimental data shows very poor agreement with the theoretically predicted behavior in equation (34).

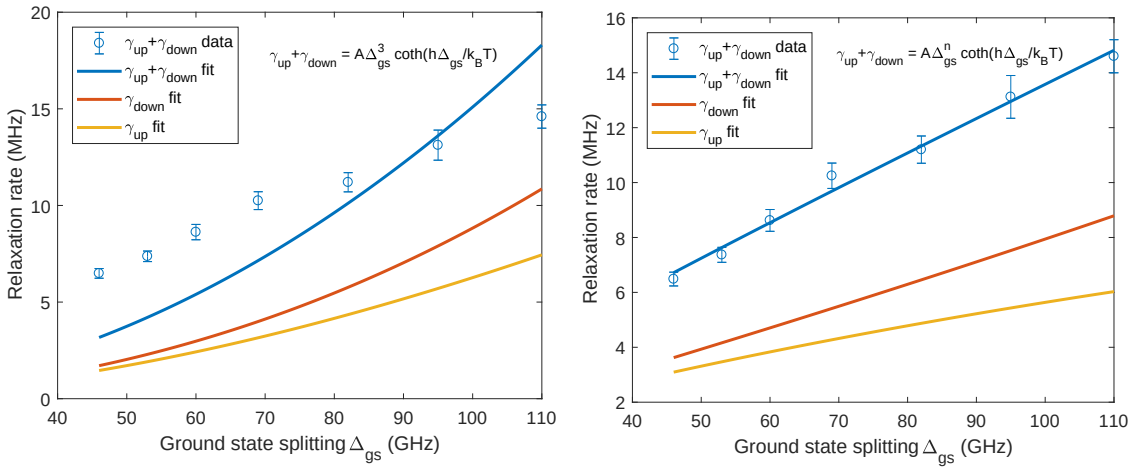

Figure S14: Fits to experimental measurements of net orbital thermalization rate $\gamma_{\mathrm{up}} + \gamma_{\mathrm{down}}$ (solid blue line) assuming coupling to only bulk acoustic modes (left plot), and unknown DOS behavior (right plot) with varying n. Fit expressions are indicated in top left corner of each plot. Corresponding individual rates γ_{up} and γ_{down} are extracted and plotted after fitting in each case (solid yellow and orange lines). The fit to variable DOS (right plot) yields $n = 1.9 \pm 0.3$.

Instead, when we allow the exponent of the DOS term to be a variable n, we obtain a better fit yielding $n = 1.9 \pm 0.3$. We can arrive at a possible explanation for this discrepancy, if we note that the SiV is situated at a nominal depth of 50 nm, which is sub-wavelength for phonons in the frequency range of Δ_{gs} probed in our experiment. As a result, we can expect appreciable coupling to surface acoustic modes, which are not accounted for in the above derivation. Since the DOS of surface modes scales as Δ_{gs}^2, it appears that the thermalization rate in our experiment is almost entirely determined by surface modes.

9 SiV spin measurement techniques

9.1 Experimental setup

The sample is cooled down to a temperature of 3.8 K inside a closed cycle liquid helium cryostat (Attodry 1000). It is placed in a dip stick, in which helium gas (pressure ~ 1 mbar) acts as an exchange gas. Two superconducting coils surrounding the sample chamber can be used to apply a magnetic field along two orthogonal axes, up to 8 T vertically and up to 2 T horizontally. DC voltage for cantilever-deflection is supplied by a high-voltage source (Stanford Research Systems PS310/1250V-25W). As an added precautionary measure described in Section 4.1, the weak leakage-current in the circuit (typically below 100 nA) is monitored via a Keithley 2400 source-meter. The optical part of the setup consists of a home-built confocal microscope mounted on top of the cryostat, and a microscope objective (NA = 0.82) inside the sample chamber. The sample is mounted on piezoelectric stages (Attocube ANPx101 and ANPz101) allowing to position the sample with respect to the objective. Non-resonant excitation of SiVs is performed using a diode laser at 660 nm (Laser Quantum Ventus), while resonant excitation is achieved with a tunable diode laser around 737 nm (Toptica DLpro). The frequency of the latter is stabilised through continuous feedback from a wavemeter (High Finesse WSU).

9.2 Coherent-population-trapping

For coherent population trapping (CPT), sidebands are generated on the resonant excitation laser using an EOM (Photline NIR-MX800) connected to a tunable microwave source (Rhode&Schwarz SMF 100A). Fluorescence from the emitters is collected through the microscope objective. A 750 nm long-pass filter in the confocal microscope allows collection of the phonon-sideband emission from SiV^- centres, filtering out the laser excitation. This emission is then sent to an avalanche photodiode (APD) (Excellitas). This measurement is repeated for different values of ground state splitting, the results of which constitute Figs. 4b and 4c of the main text.

9.3 Spin lifetime measurements

For spin lifetime (T_1) measurements, optical pulses are generated with an acousto-optic modulator (AOM, AAOptoelectronic MT350-A0.12-800) controlled by a digital delay generator (Stanford Research Instruments DG645). The signal from the APD is sent to a time-to-digital converter (qutools quTAU) triggered by the delay generator. The spin lifetime is evaluated through a pump-probe measurement at a temperature of 3.8 K and a magnetic field of 0.2 T making an angle of 54.7° with the SiV^- symmetry axis. A first 200 ns laser pulse resonant with the transition between states labelled $|1\downarrow\rangle$ and $|3\uparrow\rangle$ (Fig. 4a of the main text) initialises the spin state of the SiV^- by optical pumping into the spin up state $|1\uparrow\rangle$ (similar to initialisation of orbital population in Fig. S9). A second 200 ns pulse identical to the first one measures the population recovery into the spin down state $|1\downarrow\rangle$, and thus facilitates spin readout. By varying the time delay between the two pulses, we observe an exponential recovery in spin population determined by the T_1. This measurement is repeated

for different values of ground state splitting, the result of which is shown in Fig. 4e of the main text.

9.4 Calibration of ground and excited state splittings with applied voltage at high strain

The ground state splitting can be measured directly by considering the energy difference between the transitions labelled C and D in Fig. 2(a) of the main text. Due to the presence of other SiV^- centres generated at the same spot due to the FIB implantation procedure, measuring this energy difference between C and D transitions through non-resonant PL spectra is impractical. At the same time, simultaneous detection of C and D transitions through resonant excitation is not possible at high strain, since the Boltzmann population in level $|2\rangle$ becomes negligible as discussed in Section 7. To overcome this limitation, we resonantly excite transition A of the SiV center being studied, and record the spectrum of transitions C and D on a spectrometer after having filtered out the resonant laser with a monochromator. The spectra are then fitted with two Lorentzian functions to extract the value of the ground state splitting, as shown in Fig. S15. Likewise, the excited state splitting (also shown in the figure) can be derived as the difference between the frequencies of transitions A and C in this measurement.

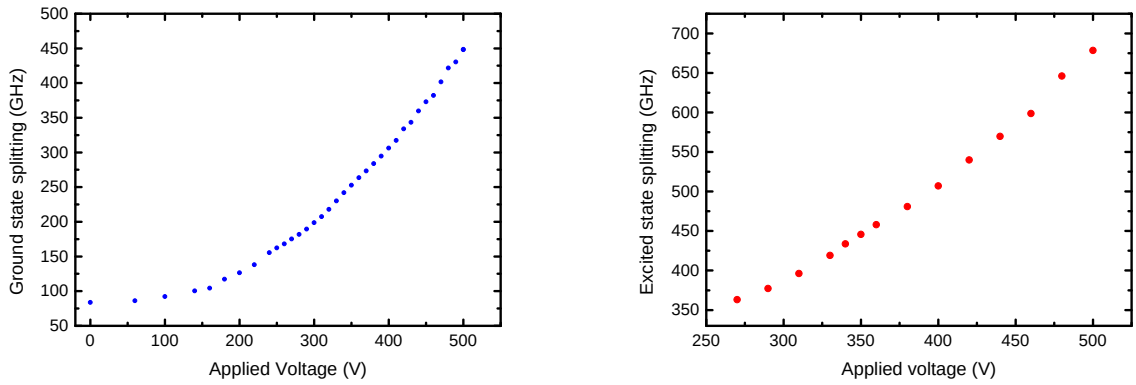

Figure S15: SiV ground state splitting (left), and excited state splitting (right) measured as a function of voltage applied to the cantilever. Error bars are smaller than the dots and correspond to the standard deviation on the frequency difference between C and D transitions estimated from Lorentzian fits.

10 Investigation of double-dip CPT signal

We first rule out the hypothesis that the two dips in our CPT measurements originate from two different SiV centres. The dips are very similar in width and depth, and their frequency-separation remains constant (4.0 ± 0.1 MHz) over a wide range of applied strain (see Fig. 4b of main text). In the event that this is caused by two SiV centers, they are required to

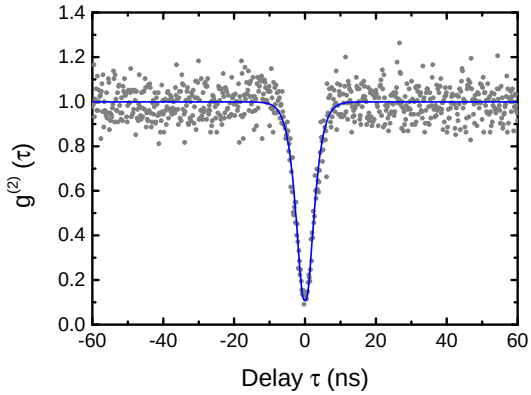

Figure S16: Second-order correlation measurement of the SiV^- centre investigated. The measured data points are plotted as grey dots. A fit based on a three-level model and accounting for timing jitter is plotted as a blue curve.

have the same fluorescence intensity, and experience exactly the same strain conditions at any applied voltage, which is very unlikely. In order to categorically establish that we are investigating a single SiV centre, we perform a Hanbury-Brown-Twiss (HBT) experiment on the phonon-sideband fluorescence upon resonant excitation of transition C, and measure the second order correlation function $g^{(2)}$ as shown in Fig. S16. The $g^{(2)}$ function is fitted using a three-level model described in [22] convolved with the Gaussian response of the avalanche photodiodes used, which have a timing jitter of 350 ps. At zero time delay, a clear anti-bunching reaching $g^{(2)}(\tau = 0) = 0.12$ indicates that the measured photons originate from a single emitter.

To gain more insight into the origin of the two CPT dips, we perform CPT at varying orientation of the applied magnetic field, while keeping its magnitude (0.2 T) constant. We work in the high-strain regime at a ground state splitting of 467 GHz. In our results, shown in Fig. S17, we observe that the separation between the two CPT dips displays a periodic variation as the magnetic field is rotated.

Given the similarity of the two dips, a very plausible explanation for the double-dip structure is the presence of a proximal spin in the environment of the SiV centre being studied. Physically, varying the direction of the applied B-field leads to a variation in the quantization axis of the SiV electron-spin (or semi-classically, the orientation of the electron magnetic moment). Likewise, the quantization axis of the proximal spin has its own variation with the B-field orientation. For instance, if this proximal spin is a nuclear spin, to first order, its orientation simply follows that of the applied B-field. As a result, the dipole-dipole interaction energy of the SiV electron-spin with the neighboring spin varies with B-field orientation, leading to the periodic behavior observed experimentally in Fig. S17. Below, we describe a semi-classical approach to model the CPT dip separation as a dipole-dipole interaction.

The Hamiltonian for two dipoles with magnetic moments $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ is given by

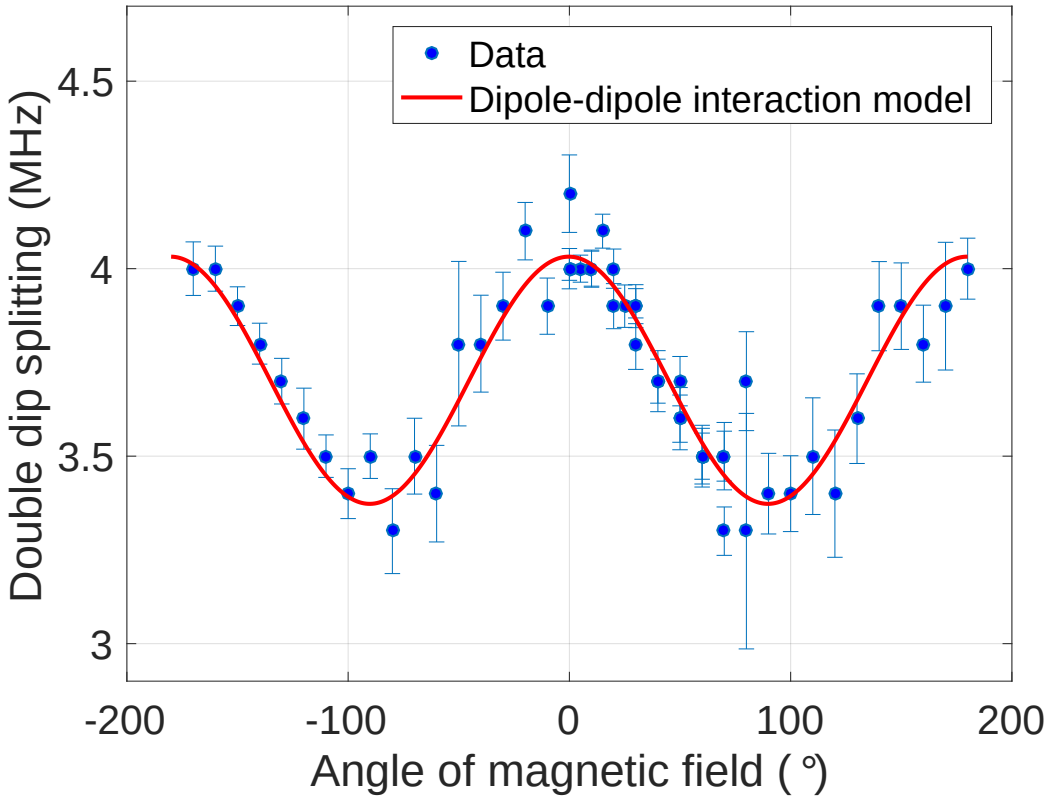

Figure S17: Dependence of CPT dip separation on magnetic field orientation. The angle plotted on the x-axis is measured with respect to the vertical direction on the sample. 0° corresponds to the [001] axis of diamond, while 90° corresponds to the [110] axis of diamond, along which the cantilever long-axis is aligned. The SiV investigated is a transverse SiV, so its internal Z-axis is either $[1\bar{1}1]$ or $[\bar{1}11]$. Error bars correspond to the standard deviation on the CPT dip frequencies estimated from Lorentzian fits.

$$\mathbb{H}^{\text{d-d}} = -\frac{\mu_0}{4\pi \left|\mathbf{r}\right|^3} \left(3 \left(\boldsymbol{\mu_1} \cdot \hat{\mathbf{r}}\right) \left(\boldsymbol{\mu_2} \cdot \hat{\mathbf{r}}\right) - \boldsymbol{\mu_1} \cdot \boldsymbol{\mu_2}\right) \tag{35}$$

where μ_0 is the vacuum-permeability, $\mathbf{r}$ is the vector from one dipole to the other, and $\hat{\mathbf{r}}$ is given by $\mathbf{r}/\left|\mathbf{r}\right|$. If the two dipoles are spins described by spin angular momentum $\mathbf{S}_1$ and $\mathbf{S}_2$, we can write the interaction Hamiltonian in terms of the spin-operators.

$$\mathbb{H}^{\text{d-d}} = -\frac{\mu_0 \gamma_1 \gamma_2}{4\pi \left|\mathbf{r}\right|^3} \left(3 \left(\mathbf{S}_1 \cdot \hat{\mathbf{r}}\right) \left(\mathbf{S}_2 \cdot \hat{\mathbf{r}}\right) - \mathbf{S_1} \cdot \mathbf{S_2}\right) \tag{36}$$

where γ_i is the gyromagnetic ratio of spin i. Semi-classically, we can treat spin angular momentum as a vector quantity $\mathbf{S} = \frac{\hbar}{2}\left(\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle\right)$, where $\sigma_x, \sigma_y, \sigma_z$ are Pauli spin-matrices. This vector describes the mean-orientation of the electron magnetic moment.

To calculate the SiV electron-spin orientation under our experimental conditions, the full ground-state Hamiltonian including spin-orbit coupling, external strain, and magnetic field must be diagonalized. The effect of the external magnetic field is described by the Zeeman Hamiltonian below written in the basis $\{|e_X,\uparrow\rangle, |e_X,\downarrow\rangle, |e_Y,\uparrow\rangle, |e_Y,\downarrow\rangle\}$,

$$\mathbb{H}^{\text{Zeeman}} = q\gamma_L \hat{L}_z B_z + \gamma_S \hat{\mathbf{S}}\cdot\mathbf{B} = q\gamma_L \begin{bmatrix} 0 & 0 & iB_z & 0 \\ 0 & 0 & 0 & iB_z \\ -iB_z & 0 & 0 & 0 \\ 0 & -iB_z & 0 & 0 \end{bmatrix} + \gamma_S \begin{bmatrix} B_z & B_x - iB_y & 0 & 0 \\ B_x + iB_y & -B_z & 0 & 0 \\ 0 & 0 & B_z & B_x - iB_y \\ 0 & 0 & B_x + iB_y & -B_z \end{bmatrix} \quad (37)$$

where the first and second terms are from the orbital angular momentum, and the spin angular momentum, respectively. $\hat{L}_z$ and $\hat{S}$ are L_z and $\mathbf{S}$ normalized by $\hbar$, respectively. The gyromagnetic ratios for each term are given by $\gamma_L = \mu_B/\hbar$, $\gamma_S = 2\mu_B/\hbar$, where μ_B is the Bohr magneton. q is a quenching factor that is commonly observed in solid-state emitters [18]. Now the total Hamiltonian is given by combining equations (17) and (37).

$$\mathbb{H}^{\text{total}} = \mathbb{H}^{\text{SO}} + \mathbb{H}^{\text{strain}} + \mathbb{H}^{\text{Zeeman}} \quad (38)$$

We diagonalize the total Hamiltonian, $\mathbb{H}^{\text{total}}$, and calculate expectation values of the Pauli matrices for the lowest two eigenstates, which comprise the SiV spin-qubit under investigation. This gives us the mean orientation of the SiV electron-spin, say $\mathbf{S}_1$ for given experimental conditions. To calculate the mean orientation of the proximal spin $\mathbf{S}_2$, we assume it to be either a nuclear spin such as ^{13}C, or an electron spin such as another SiV-center. In the case of a nuclear spin, $\mathbf{S}_2$ is simply given by the direction of the external magnetic field. Once the quantization axes of the two spins are known, we can fit our data to the calculated value of $\mathbb{H}^{\text{d-d}}$ from Eq. 36 by using the distance between the spins $\mathbf{r}$ as a fit parameter. The result of such a fitting procedure is shown in Fig. S17. In the case of a nuclear spin, the distance between the two spins $|\mathbf{r}|$ is estimated to be on the order of 1 Å. If the other spin is an electron spin from another SiV centre, it is possible to obtain similar results as in Fig. S17. However, in this case, the distance between the spins is on the order of tens of nanometres.

11 Spectral diffusion of C transition

The C transition of the SiV centre used in our CPT measurements does not experience extra spectral diffusion at high strain (high voltage). In theory, there are two potential causes of spectral diffusion that might appear at high strain. (i) A gradient of strain could perturb the inversion symmetry of the SiV-center, and induce a nonzero permanent dipole moment for the electronic levels. (ii) A large DC-electric field at the SiV location arising from the applied voltage could induce a linear Stark shift response to fluctuating electric-fields

in the environment of the SiV, thereby increasing spectral diffusion. In our experiments, the amount of spectral diffusion on the timescale of minutes (roughly ±20 MHz) does not appear to increase from the zero-voltage to the high-voltage condition (Fig. S18). This is further confirmed by our measurements of the C transition linewidth at various high strain conditions. (Fig. S19). We do observe a systematic drift in the transition frequency on the timescale of hours, but this slow drift can be easily corrected for by adjusting the resonant laser excitation frequency (or alternatively by performing feedback on the DC voltage applied to our device).

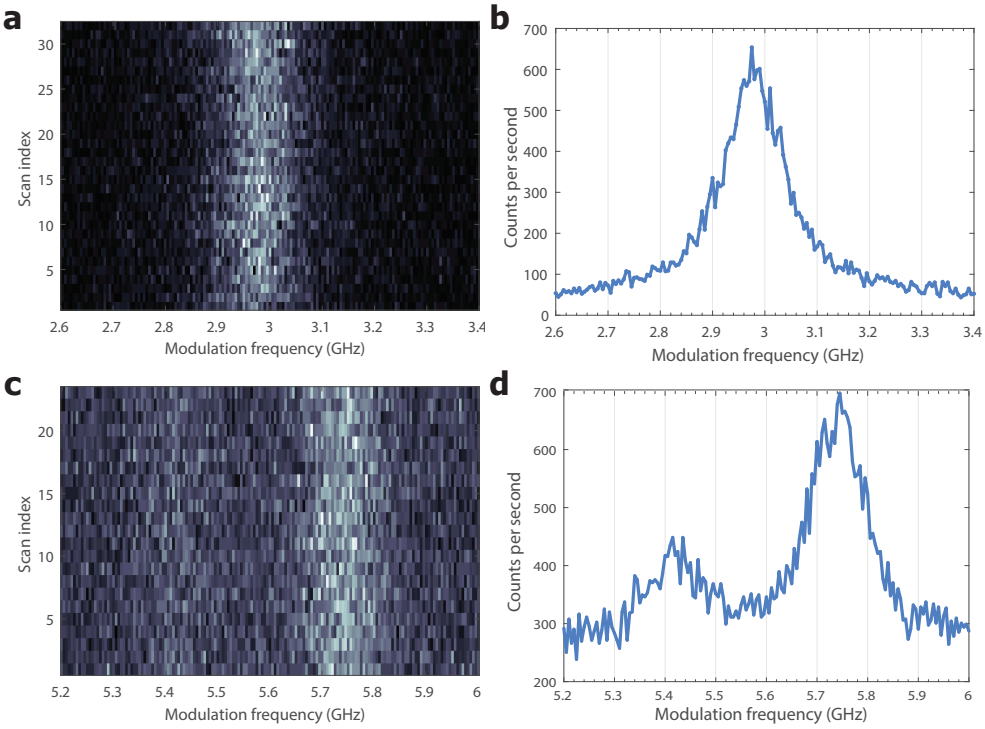

Figure S18: (a) Continuous monitoring of the resonant excitation spectrum of the C transition at 0 V (ground state splitting of 88 GHz) and 0 T. Each horizontal line corresponds to a single acquisition of the spectrum over a duration of 0.6 s with 5 MHz spectral resolution. (b) Average of the 32 scans in part (a) corresponding to a total acquisition time of 20 s (c) Continuous monitoring of the resonant excitation spectrum of two of the four Zeeman-split C transitions at 400 V (ground state splitting of 303 GHz) and a B-field of 0.2 T. Each horizontal line corresponds to a single acquisition of the spectrum over a duration of 0.6 s with 5 MHz spectral resolution. (d) Average of the 23 scans in part (c) corresponding to a total acquisition time of 14 s.

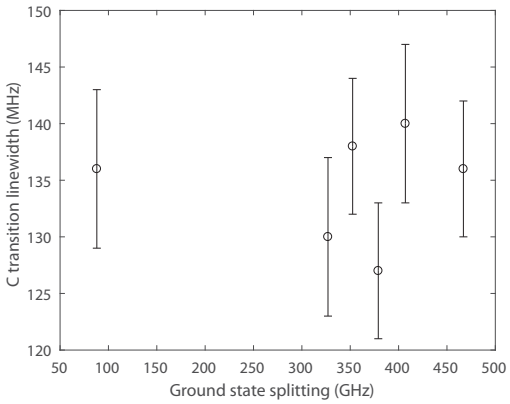

Figure S19: Linewidth of the C transition measured at various strain conditions over a timescale of few minutes. Error bars are obtained from Lorentzian fits to the spectra.

12 Spin relaxation (T_1) model

E_g-phonons predominantly drive spin-conserving transitions between the GS orbitals of the SiV (i.e. between $\{|1\downarrow\rangle, |2\downarrow\rangle\}$, and $\{|1\uparrow\rangle, |2\uparrow\rangle\}$ respectively). These processes affect spin T_2^*, but are expected to preserve spin T_1. However, in the presence of an off-axis magnetic-field, and non-zero lattice strain, the eigenstates of the GS manifold are no longer pure spin-orbit eigenstates, and all transitions between the four states within the GS manifold become allowed for E_g-phonons. Here we consider various channels for phonon-mediated spin-flips, and develop a microscopic model for spin-relaxation of the qubit defined by $|1\downarrow\rangle, |1\uparrow\rangle$. Towards the end of this section, we conclude that spin-relaxation at 4K is dominated by a two-phonon process involving the upper ground state orbital branches as intermediate states. In literature, this is sometimes referred to as an Orbach process after [23]. The experimentally observed behavior of spin T_1 increasing with the GS splitting Δ_{gs} in Fig. 3e of the main text is well-explained by such a process.

Specifically, consider the case in which spin-population in the lower orbital-branch $|1\rangle$ is initially polarized into the lower spin-sublevel $|1\downarrow\rangle$. We consider the following pathways for population to relax into the upper spin-sublevel $|1\uparrow\rangle$:

- Direct single-phonon relaxation: Via a single phonon of frequency ω_{s} resonant with the spin-transition as shown in Fig. S20(a)
- Resonant two-phonon relaxation: Via two phonons resonant with a level in the upper orbital branch as an intermediate state as shown in Fig. S20(b). The spin-flip can be caused by either the emitted phonon (left) or the absorbed phonon (right).
- Off-resonant two-phonon relaxation: Via two phonons with a virtual level as an intermediate state as shown in Fig. S20(c). The effective driving strength of each phonon will be reduced from its value in the resonant process by an amount corresponding to the detuning from the upper orbital branch.

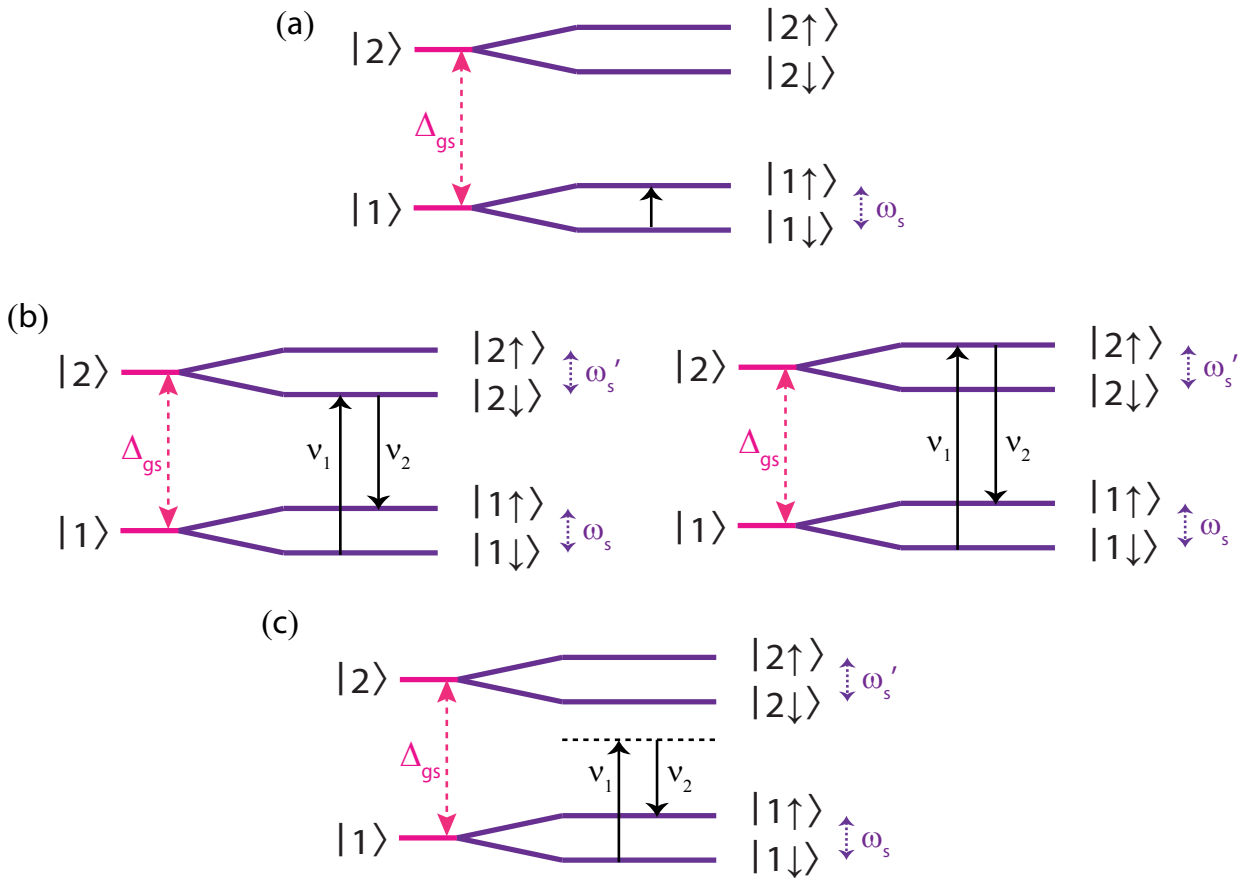

Figure S20: Various pathways for a phonon-mediated spin-flip (a) Direct relaxation via a single phonon resonant with the $|1\downarrow\rangle \rightarrow |1\uparrow\rangle$ spin-transition. (b) Two possible channels for a resonant two-phonon process involving the upper orbital branch. (c) Off-resonant two-phonon processes.

12.1 Single-phonon spin-relaxation

The relaxation rate due to single-phonon absorption at the frequency of the spin qubit, ω_s can be calculated using Fermi's golden rule. We will obtain a result similar to to 32.

$$\Gamma_{\text{up,dir}} = 2\pi\chi_s\rho\omega_s^3 n_{th}(\omega_s) \tag{39}$$

Importantly, the quantity χ_s determines the mean-squared single-phonon coupling strength and is related to χ for the orbital transition in equation 31 by

$$\chi_s = \left(\frac{d_{1\downarrow\rightarrow 1\uparrow}}{d}\right)^2 \chi \tag{40}$$

Here $d_{1\downarrow\rightarrow 1\uparrow}$ is the strain-susceptibility for the spin-flipping $|1\downarrow\rangle \rightarrow |1\uparrow\rangle$ transition, and is the only unknown, while d (see table 2) is the experimentally known strain-susceptibility

of the GS-orbitals. We can estimate $d_{1\downarrow\to 1\uparrow}$ in terms of d by considering the GS Hamiltonian.

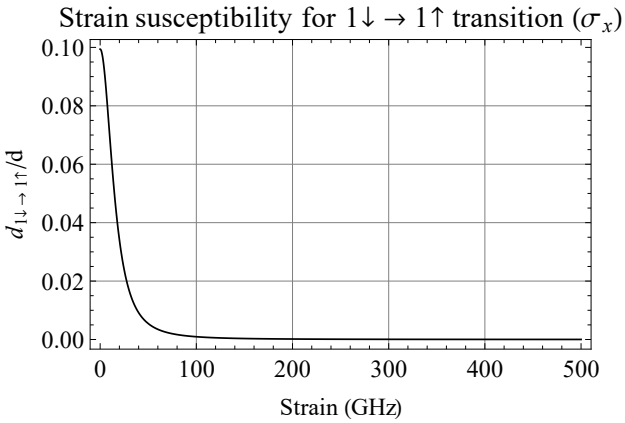

Figure S21: Strain-susceptibility for the $1\downarrow\to 1\uparrow$ transition (in units of d) vs. strain β. B-field of 0.2 T aligned with the $\langle 001\rangle$ axis.

In particular, at zero strain and zero B-field, the eigenstates of the GS Hamiltonian $\{|1\downarrow\rangle, |1\uparrow\rangle, |2\downarrow\rangle, |2\uparrow\rangle\}$ are the spin-orbit eigenstates $\{|e_+\downarrow\rangle, |e_-\uparrow\rangle, |e_-\downarrow\rangle, |e_+\uparrow\rangle\}$, where $|e_\pm\downarrow\rangle = |e_X\downarrow\rangle \pm |e_Y\downarrow\rangle$, and $|e_\pm\uparrow\rangle = |e_X\uparrow\rangle \pm |e_Y\uparrow\rangle$ [18] [1]. Now, the GS Hamiltonian in 38 can be written in the basis of these spin-orbit eigenstates. This basis-switch gives us the following constituent spin-orbit, Zeeman and strain Hamiltonians.

$$\mathbb{H}^{\text{SO}} = \begin{bmatrix} -\frac{\lambda_{SO}}{2} & 0 & 0 & 0 \\ 0 & -\frac{\lambda_{SO}}{2} & 0 & 0 \\ 0 & 0 & \frac{\lambda_{SO}}{2} & 0 \\ 0 & 0 & 0 & \frac{\lambda_{SO}}{2} \end{bmatrix} \tag{41}$$

$$\mathbb{H}^{\text{Zeeman}} = \begin{bmatrix} -\gamma_L B_z - \gamma_s B_z & 0 & 0 & \gamma_s B_x \\ 0 & \gamma_L B_z + \gamma_s B_z & \gamma_s B_x & 0 \\ 0 & \gamma_s B_x & \gamma_L B_z - \gamma_s B_z & 0 \\ \gamma_s B_x & 0 & 0 & -\gamma_L B_z + \gamma_s B_z \end{bmatrix} \tag{42}$$

$$\mathbb{H}^{\text{strain}} = \begin{bmatrix} 0 & 0 & \beta + i\gamma & 0 \\ 0 & 0 & 0 & \beta + i\gamma \\ \beta - i\gamma & 0 & 0 & 0 \\ 0 & \beta - i\gamma & 0 & 0 \end{bmatrix} \tag{43}$$

Without loss of generality, we have assumed $B_y = 0$, and also set the α term in $\mathbb{H}^{\text{strain}}$ to zero, since it only describes a common mode shift. Now further, assume $\gamma = 0$, and that

[1] The notation $\{|1\downarrow\rangle, |1\uparrow\rangle, |2\downarrow\rangle, |2\uparrow\rangle\}$ is simplistic. Note that at zero-strain, and zero B-field, the orbital characters of $|1\downarrow\rangle, |1\uparrow\rangle$ (and of $|2\downarrow\rangle, |2\uparrow\rangle$) are orthogonal. As strain is applied, and increased far above the SO-coupling λ_{SO}, the orbital characters of $|1\downarrow\rangle, |1\uparrow\rangle$ (and of $|2\downarrow\rangle, |2\uparrow\rangle$) tend towards becoming identical. Further, as discussed in Section 11, the spin-quantization axis is determined by a combination of SO-coupling, external strain, and direction of external B-field.

the strain at the SiV site is entirely in the β-mode (this is fairly accurate for transverse SiVs as can be inferred from Fig. S3c and equation (14)). We can now diagonalize the total Hamiltonian $\mathbb{H} = \mathbb{H}^{\text{SO}} + \mathbb{H}^{\text{Zeeman}} + \mathbb{H}^{\text{strain}}$ to obtain the eigenstates $\{|1\downarrow\rangle, |1\uparrow\rangle, |2\downarrow\rangle, |2\uparrow\rangle\}$ in terms of the spin-orbit eigenstates. Then $d_{1\downarrow\to 1\uparrow}$ is given by

$$d_{1\downarrow\to 1\uparrow} = \frac{\langle 1\uparrow|\mathbb{H}^{\text{strain}}|1\downarrow\rangle}{\langle e_+\downarrow|\mathbb{H}^{\text{strain}}|e_-\downarrow\rangle} d = \frac{\langle 1\uparrow|\mathbb{H}^{\text{strain}}|1\downarrow\rangle}{\beta} d \tag{44}$$

Fig. S21 shows a plot of $d_{1\downarrow\to 1\uparrow}$ as a function of the DC-strain β, while assuming a B-field of 0.2 T aligned with the [001] axis. In fact, at low strain, we can use perturbation theory to arrive at $d_{1\downarrow\to 1\uparrow} \approx \frac{\gamma_s B_x}{\Delta_{\text{gs}}} d$ ($\approx \frac{\gamma_s B_x}{\lambda_{\text{SO}}} d$ at zero strain). We see that this behavior agrees with the result of the exact calculation in Fig. S21 the low-strain regime.

With $d_{1\downarrow\to 1\uparrow}$ known, we can calculate $\Gamma_{\text{up,dir}}$ according to equation (39). In fact, combining equations 32, 39, 40, we can write $\Gamma_{\text{up,dir}}$ in terms of the upward orbital relaxation rate at zero-strain, $\gamma_{\text{up}}(46\text{ GHz})$ as follows -

$$\Gamma_{\text{up,dir}} = \left(\frac{\omega_s}{46\text{GHz}}\right)^3 \frac{n_{th}(\omega_s)}{n_{th}(46\text{GHz})} \left(\frac{d_{1\downarrow\to 1\uparrow}}{d}\right)^2 \gamma_{\text{up}}(46\text{GHz}) \tag{45}$$

Similarly, we have

$$\Gamma_{\text{down,dir}} = \left(\frac{\omega_s}{46\text{GHz}}\right)^3 \frac{n_{th}(\omega_s)+1}{n_{th}(46\text{GHz})+1} \left(\frac{d_{1\downarrow\to 1\uparrow}}{d}\right)^2 \gamma_{\text{down}}(46\text{GHz}) \tag{46}$$

Then the total direct single-phonon relaxation rate is given by

$$\Gamma_{\text{dir}} = \Gamma_{\text{up,dir}} + \Gamma_{\text{down,dir}} \tag{47}$$

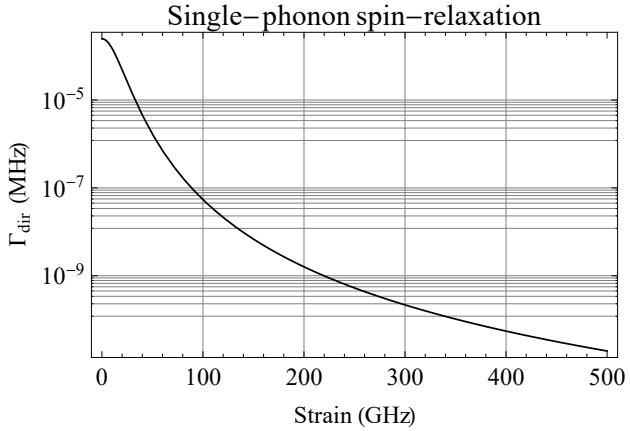

Figure S22: Direct single-phonon spin-relaxation rate vs. strain β. (B-field of 0.2 T aligned with the [001] axis, and T =4 K)

With $\gamma_{\text{up}}(46\text{GHz}) + \gamma_{\text{down}}(46\text{GHz}) = 4$ MHz, $\frac{\gamma_{\text{up}}(46\text{GHz})}{\gamma_{\text{down}}(46\text{GHz})} = \frac{n_{th}(46\text{GHz})}{n_{th}(46\text{GHz})+1}$, and assuming the model for $\gamma_{\text{up}}(\Delta_{\text{gs}})$ developed in Section 9, we plot Γ_{dir} against strain in Fig. S22. The calculated value at zero strain is 3-4 orders of magnitude weaker than the orbital relaxation

rate, due to the weaker strain-susceptibility, and lower DOS of phonons driving the spin transition. At high strain, Γ_{dir} is expected to decay as $\sim 1/\Delta_{\text{gs}}^2$ following the variation in $d^2_{1\downarrow\to 1\uparrow}$.

12.2 Resonant two-phonon spin-relaxation

The two pathways for resonant two-phonon relaxation shown in Fig. S20(b) are driven by phonons of different frequencies; the values of ν_1, ν_2 are different between the processes in the left and right panels. Since we do not expect any phase-coherence between these phonons, we can treat each pathway independently. Consider first the pathway $1\downarrow\to 2\uparrow\to 1\downarrow$ as shown in the right panel of Fig. S20(b). We can write the effective two-phonon transition rate from $1\downarrow$ to $1\uparrow$ as

$$\tilde{g}_{1\downarrow\to 2\uparrow\to 1\uparrow}(\mathbf{k_1},\lambda_1;\mathbf{k_2},\lambda_2) = \frac{g_{1\downarrow\to 2\uparrow}(\mathbf{k_1},\lambda_1)g_{2\uparrow\to 1\uparrow}(\mathbf{k_2},\lambda_2)}{\gamma_{\text{down}}} \tag{48}$$

Here $\mathbf{k_i}, \lambda_i$ are the wave-vector and polarization corresponding to the phonon at frequency ν_i in Fig. S20(b). g_P indicates a transition rate for pathway P. We can write

$$\tilde{g}_{1\downarrow\to 2\uparrow\to 1\uparrow}(\mathbf{k_1},\lambda_1;\mathbf{k_2},\lambda_2) = \frac{\frac{d_{1\downarrow\to 2\uparrow}}{d} g(\mathbf{k_1},\lambda_1)g(\mathbf{k_2},\lambda_2)}{\gamma_{\text{down}}} = \frac{d_{1\downarrow\to 2\uparrow}}{d}\frac{g(\mathbf{k_1},\lambda_1)g(\mathbf{k_2},\lambda_2)}{\gamma_{\text{down}}}$$

Here g is the transition rate corresponding to a spin-conserving orbital transition as defined by equation 30. The relaxation rate through this pathway is thus given by

$$\begin{aligned}\Gamma_{1\downarrow\to 2\uparrow\to 1\uparrow} = 2\pi \sum_{\mathbf{k_1},\lambda_1}\sum_{\mathbf{k_2},\lambda_2} & \left|\frac{d_{1\downarrow\to 2\uparrow}}{d}\frac{g(\mathbf{k_1},\lambda_1)g(\mathbf{k_2},\lambda_2)}{\gamma_{\text{down}}}\right|^2 n_{th}(\nu_1)\left[n_{th}(\nu_2)+1\right] \\ & \delta\left(\nu_1 - \Delta_{\text{gs}} - \frac{\omega_s}{2} - \frac{\omega_s'}{2}\right)\delta\left(\nu_2 - \nu_1 + \omega_s\right)\end{aligned} \tag{49}$$

At strain $\beta \gg \lambda_{\text{SO}}$, we have $\omega_s \approx \omega_s'$ allowing us to set the conditions within the first delta function to $\nu_1 - \Delta_{\text{gs}} - \omega_s$. Converting this sum into an integral, we get

$$\begin{aligned}\Gamma_{1\downarrow\to 2\uparrow\to 1\uparrow} = 2\pi \int & \left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \frac{\chi\nu_1\cdot\chi(\nu_1-\omega_s)}{\gamma_{\text{down}}^2} n_{th}(\nu_1)\left[n_{th}(\nu_1-\omega_s)+1\right] \\ & \rho\nu_1^2\cdot\rho(\nu_1-\omega_s)^2 2\pi\gamma_{\text{down}}\delta(\nu_1-\Delta_{\text{gs}}-\omega_s)d\nu_1\end{aligned} \tag{50}$$

Here χ and ρ are the same quantities defined for orbital transitions in Section 9. Finally, the condition $\omega_s \ll \Delta_{\text{gs}}$ gives us

$$\Gamma_{1\downarrow\to 2\uparrow\to 1\uparrow} = 4\pi^2\left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \frac{\chi^2\rho^2\Delta_{\text{gs}}^6}{\gamma_{\text{down}}} n_{th}(\Delta_{\text{gs}}) \tag{51}$$

Using equations 32,33 at $h\Delta_{\text{gs}} \gg k_B T$, this simplifies to

$$\Gamma_{1\downarrow\to 2\downarrow\to 1\uparrow} = \left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \gamma_{\text{up}} \tag{52}$$

Examining the RHS, we conclude that this rate simply corresponds to the single-phonon absorption rate along the path $1\downarrow \rightarrow 2\uparrow$, the first leg of the two-phonon process. Intuitively, this leg involves a spin-flip in the already slower absorption process, and is thus expected to be the 'rate-determining step'. We can analyze the other two-phonon relaxation pathway $1\downarrow \rightarrow 2\downarrow \rightarrow 1\uparrow$ shown in the left panel of Fig. S20(b), and arrive at the same result. Finally, including processes in the opposite direction from $1\uparrow \rightarrow 1\downarrow$, the total resonant two-phonon relaxation rate can be written as

$$\Gamma_{\text{2ph,res}} = 4\Gamma_{1\downarrow\rightarrow 2\uparrow\rightarrow 1\uparrow} = 16\pi^2 \left(\frac{d_{1\downarrow\rightarrow 2\uparrow}}{d}\right)^2 \frac{\chi^2\rho^2\Delta_{\text{gs}}^6}{\gamma_{\text{down}}} n_{th}(\Delta_{\text{gs}}) \tag{53}$$

$$= 4\left(\frac{d_{1\downarrow\rightarrow 2\uparrow}}{d}\right)^2 \gamma_{\text{up}} \quad \text{for} \quad h\Delta_{\text{gs}} \gg k_B T \tag{54}$$

Similar to equation 44, we can write

$$d_{1\downarrow\rightarrow 2\uparrow} = \frac{\langle 2\uparrow|\mathbb{H}^{\text{strain}}|1\downarrow\rangle}{\beta} d \tag{55}$$

$$d_{2\downarrow\rightarrow 1\uparrow} = \frac{\langle 2\downarrow|\mathbb{H}^{\text{strain}}|1\uparrow\rangle}{\beta} d \tag{56}$$

Fig. S23 shows a plot of $d_{1\downarrow\rightarrow 2\uparrow}$, and $d_{2\downarrow\rightarrow 1\uparrow}$ as a function of the DC-strain β, while assuming a B-field of 0.2 T aligned with the [001] axis. They are zero at zero strain, rapidly increase to a significant value of $\approx 0.5d$, when $\beta = \lambda_{SO}/2 = 23$ GHz. This corresponds to maximal mixing between states having opposite orbital- and opposite spin-character (i.e. $|e_+\downarrow\rangle$ mixing with $|e_-\uparrow\rangle$, $|e_-\downarrow\rangle$ mixing with $|e_+\uparrow\rangle$, and vice-versa). At even higher strain, the susceptibilities fall off as $\sim 1/\Delta_{\text{gs}}$. Fig. S24 plots the relaxation rate predicted by equation 54 assuming the same conditions on the zero strain relaxation rates as used for Fig. S22. We observe good quantitative agreement with the measured rates in Fig. 3e of the main text, suggesting that this is most likely the dominant spin-relaxation mechanism.

Since $d_{1\downarrow\rightarrow 2\uparrow}/d \sim \Delta_{\text{gs}}^{-1}$, following equation 53, we fit our experimental data in Fig. 3e to

$$\Gamma_{\text{2ph,res}} = \frac{A\Delta_{\text{gs}}}{\exp(h\Delta_{\text{gs}}/k_B T) - 1} \tag{57}$$

The two-parameter fit yields $A = 0.016 \pm 0.003$, and $T = 4.3 \pm 0.3$K.

12.3 Off-resonant two-phonon spin-relaxation

Similar to equation 48 for the resonant case, we can write the effective two-phonon transition rate for the Raman process in Fig. S20(c). We will have to consider Raman transitions mediated by interaction with the levels $|2\downarrow\rangle$ and $|2\uparrow\rangle$ simultaneously. Assuming $\omega_s \approx \omega_s'$, and setting $\nu = (\nu_1 + \nu_2)/2$ gives us

$$\tilde{g}_{1\downarrow\rightarrow v\rightarrow 1\uparrow} = \frac{g_{1\downarrow\rightarrow 2\downarrow}(\mathbf{k_1},\lambda_1) g_{2\downarrow\rightarrow 1\uparrow}(\mathbf{k_1},\lambda_1)}{\Delta_{\text{gs}} - \omega_s - \nu} + \frac{g_{1\downarrow\rightarrow 2\uparrow}(\mathbf{k_1},\lambda_1) g_{2\uparrow\rightarrow 1\uparrow}(\mathbf{k_1},\lambda_1)}{\Delta_{\text{gs}} + \omega_s - \nu} \tag{58}$$

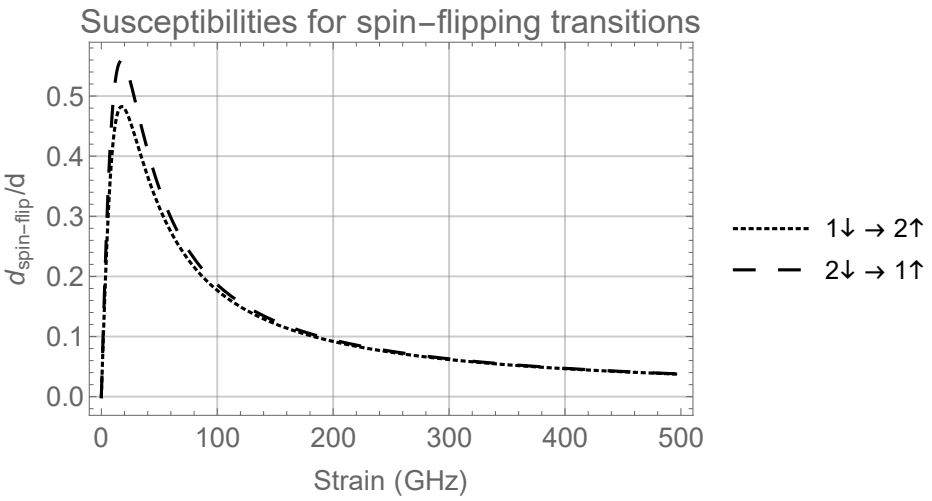

Figure S23: Strain-susceptibilities for spin-flipping transitions $1\downarrow \rightarrow 2\uparrow$ and $2\downarrow \rightarrow 1\uparrow$ (in units of d) vs. strain β. B-field of 0.2 T aligned with the $\langle 001 \rangle$ axis.

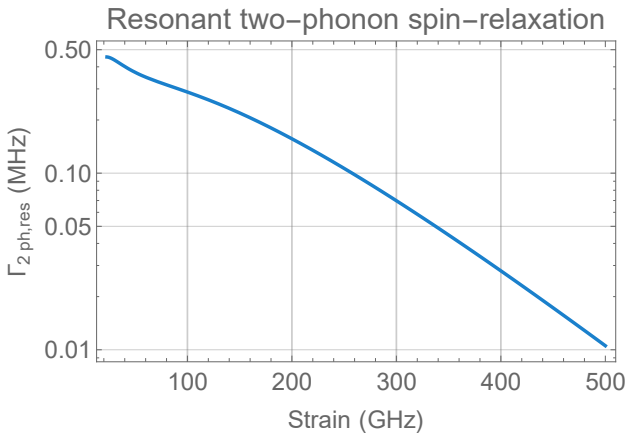

Figure S24: Resonant two-phonon spin relaxation rate vs. strain β varying from 23 GHz (maximum value for spin-flipping strain susceptibilities in equations 55,56) to 500 GHz. (B-field of 0.2 T aligned with the [001] axis, and T =4 K)

Using Kramers' theorem for a spin-1/2 system [23], we can show that

$$g_{1\downarrow\rightarrow 2\uparrow}(\mathbf{k_1},\lambda_1)g_{2\uparrow\rightarrow 1\uparrow}(\mathbf{k_1},\lambda_1) = -g_{1\downarrow\rightarrow 2\downarrow}(\mathbf{k_1},\lambda_1)g_{2\downarrow\rightarrow 1\uparrow}(\mathbf{k_1},\lambda_1) \tag{59}$$

This leads to a near cancellation of the two terms in equation 58, and gives the following quantity smaller than either term by the factor $\approx 2\omega_s/(\Delta_{\text{gs}} - \nu)$

$$\tilde{g}_{1\downarrow\rightarrow v\rightarrow 1\uparrow} = \frac{2\omega_s}{(\Delta_{\text{gs}} - \nu)^2 - \omega_s^2} g_{1\downarrow\rightarrow 2\downarrow}(\mathbf{k_1},\lambda_1)g_{2\downarrow\rightarrow 1\uparrow}(\mathbf{k_1},\lambda_1) \tag{60}$$

As in the resonant case, we sum over all phonon modes such that $\nu_2 - \nu_1 = \omega_s$, but instead, here we forego the resonance condition with the upper orbital branch, and obtain an integral similar to equation 50.

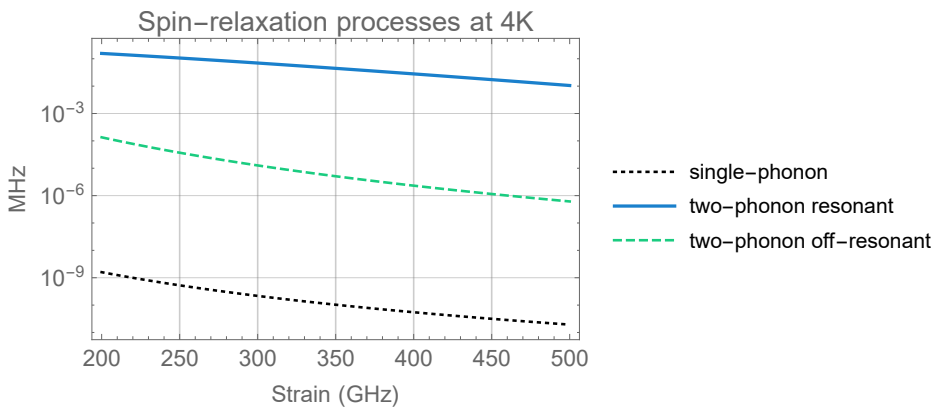

Figure S25: Rates of all three spin-relaxation mechanisms indicating their magnitudes and scaling with strain β.

$$\Gamma_{1\uparrow\to 1\downarrow,\text{off-res}} = 2\pi \int \left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \chi\nu_1 \cdot \chi(\nu_1-\omega_s)\frac{4\omega_s^2}{[(\Delta_{\text{gs}}-\nu)^2-\omega_s^2]^2} \\ n_{th}(\nu_1)\left[n_{th}(\nu_1-\omega_s)+1\right]\rho\nu_1^2\cdot\rho(\nu_1-\omega_s)^2 d\nu_1 \tag{61}$$

We assume that frequencies $\nu \sim k_B T/h$ dominate the contribution to the rate (as we saw in section 9, the product of the Bose distribution and DOS is maximized around this value). At $T = 4$ K, in our experiments, we have $\omega_s \ll k_B T \ll \Delta_{\text{gs}} \ll \Omega_D$, where Ω_D is the Debye frequency of diamond. Under these conditions, the above integral simplifies to

$$\begin{aligned}\Gamma_{1\uparrow\to 1\downarrow,\text{off-res}} &= 2\pi\left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \chi^2\rho^2\frac{4\omega_s^2}{\Delta_{\text{gs}}^4}\int_0^\infty n_{th}(\nu_1)\left[n_{th}(\nu_1)+1\right]\nu_1^6 d\nu_1 \\ &= \frac{8\pi^7 7!}{945}\left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \chi^2\rho^2\frac{\omega_s^2}{\Delta_{\text{gs}}^4}\left(\frac{k_B T}{h}\right)^7\end{aligned} \tag{62}$$

We see in Fig. S25, that at 4 K, this rate is negligible compared to the resonant two-phonon process. However, at higher temperatures $k_B T \gg h\Delta_{\text{gs}}$, this process can indeed dominate the resonant process. In this regime, the integral in 61 simplifies differently, and yields

$$\Gamma_{1\uparrow\to 1\downarrow,\text{off-res}} = 8\pi^3\left(\frac{d_{1\downarrow\to 2\uparrow}}{d}\right)^2 \chi^2\rho^2\omega_s^2\left(\frac{k_B T}{h}\right)^3 \tag{63}$$

Table 3 summarizes the discussion in this section.

Mechanism	Final result	Relevant regime	Expected scaling of rate
Single-phonon	39	$k_B T/h \ll \omega_s$	$B_\perp^2\Delta_{\text{gs}}^{-2}\omega_s^3\exp(-h\omega_s/k_B T)$
Resonant two-phonon	53	$k_B T/h \sim \Delta_{\text{gs}}$	$B_\perp^2\Delta_{\text{gs}}[\exp(h\Delta_{\text{gs}}/k_B T)-1]^{-1}$
Off-resonant two-phonon	63	$k_B T/h \gg \Delta_{\text{gs}}$	$B_\perp^2\Delta_{\text{gs}}^{-2}\omega_s^2 T^3$

Table 3: Summary of spin-relaxation mechanisms

References

[1] element six ™ (2017). URL http://www.e6.com.

[2] Hausmann, B. J. M. *et al.* Fabrication of diamond nanowires for quantum information processing applications. *Diamond & Related Materials* **19**, 621–629 (2010).

[3] Atikian, H. A. *et al.* Wafer-scale fabrication of freestanding nanostructures via reactive ion beam undercut etching (2017).

[4] Burek, M. J. *et al.* Free-Standing Mechanical and Photonic Nanostructures in Single-Crystal Diamond. *Nano Letters* **12**, 6084–6089 (2012).

[5] Sipahigil, A. *et al.* Supplementary materials - an integrated diamond nanophotonics platform for quantum-optical networks. *Science* **354**, 847–850 (2016).

[6] Schröder, T. *et al.* Scalable focused ion beam creation of nearly lifetime-limited single quantum emitters in diamond nanostructures. *arXiv:1610.09492 [quant-ph]* (2016).

[7] Chu, Y. *et al.* Coherent optical transitions in implanted nitrogen vacancy centers. *Nano Letters* **14**, 1982–1986 (2014).

[8] Evans, R. E., Sipahigil, A., Sukachev, D. D., Zibrov, A. S. & Lukin, M. D. Narrow-linewidth homogeneous optical emitters in diamond nanostructures via silicon ion implantation. *Phys. Rev. Applied* **5**, 044010 (2016).

[9] Maier, F., Riedel, M., Mantel, B., Ristein, J. & Ley, L. Origin of surface conductivity in diamond. *Phys. Rev. Lett.* **85**, 3472–3475 (2000).

[10] Chen, K. C.-Y. *et al.* Conductor fusing and gapping for bond wires. *Progress In Electromagnetics Research M* **31**, 199–214 (2013).

[11] Comsol multiphysics® v. 5.2. URL www.comsol.com.

[12] Zhan, S., Azarian, M. H. & Pecht, M. G. Surface insulation resistance of conformally coated printed circuit boards processed with no-clean flux. *IEEE transactions on electronics packaging manufacturing* **29**, 217–223 (2006).

[13] Townsend, J. S. (Oxford University Press, 1915).

[14] Strong, F. W., Skinner, J. L., Dentinger, P. M. & Tien, N. C. Electrical breakdown across micron scale gaps in mems structures. In *MOEMS-MEMS 2006 Micro and Nanofabrication*, 611103–611103 (International Society for Optics and Photonics, 2006).

[15] Sipahigil, A. *et al.* Indistinguishable Photons from Separated Silicon-Vacancy Centers in Diamond. *Physical Review Letters* **113**, 113602 (2014).

[16] Londero, E. *et al.* Ab initio study on the Herzberg-Teller effect in the optical excitation spectrum of silicon-vacancy center in diamond. *arxiv.orgabs.* (2016). 1605.02955.

[17] Hughes, A. E. & Runciman, W. A. Uniaxial stress splitting of doubly degenerate states of tetragonal and trigonal centres in cubic crystals. *Proceedings of the Physical Society* **90**, 827 (1967).

[18] Hepp, C. *Electronic Structure of the Silicon Vacancy Color Center in Diamond*. Ph.D. thesis (2014).

[19] Materials Design, I. Elastic properties of diamond. Tech. Rep. (2017).

[20] Sternschulte, H., Thonke, K., Sauer, R., Münzinger, P. C. & Michler, P. 1.681-eV luminescence center in chemical-vapor-deposited homoepitaxial diamond films. *Phys. Rev. B* **50**, 14554–14560 (1994).

[21] Jahnke, K. D. *et al.* Electron–phonon processes of the silicon-vacancy centre in diamond. *New Journal of Physics* **17**, 1–11 (2015).

[22] Neu, E. *et al.* Single photon emission from silicon-vacancy colour centres in chemical vapour deposition nano-diamonds on iridium. *New Journal of Physics* **13**, 025012 (2011).

[23] Orbach, R. Spin-lattice relaxation in rare-earth salts. *Proc. R. Soc. A* **A 264**, 458 (1961).