Supplemental Material for "Isolated spin qubits in SiC with a high-fidelity infrared spin-to-photon interface"

David J. Christle,[1] Paul V. Klimov,[1] Charles F. de las Casas,[1] Krisztián Szász,[2] Viktor Ivády,[2,3] Valdas Jokubavicius,[3] Jawad Ul Hassan,[3] Mikael Syväjärvi,[3] William F. Koehl,[1] Takeshi Ohshima,[4] Nguyen T. Son,[3] Erik Janzén,[3] Ádám Gali,[2,5] and David D. Awschalom[1,*]

[1] *Institute for Molecular Engineering, University of Chicago, Chicago, IL 60637, USA*

[2] *Wigner Research Centre for Physics, Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, PO Box 49, H-1525, Budapest, Hungary*

[3] *Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden*

[4] *National Institutes for Quantum and Radiological Science and Technology, 1233 Watanuki, Takasaki, Gunma 370-1292, Japan*

[5] *Department of Atomic Physics, Budapest University of Technology and Economics, Budafoki út 8., H-1111, Budapest, Hungary*

(Dated: June 12, 2017)

S1. APPARATUS AND METHODS

The 3C-SiC samples are grown on an off-oriented 4H-SiC substrate using sublimation epitaxial growth [1]. We use a two-step process and geometrical control approach to grow thick ($\approx 1.5\,\mathrm{mm}$), single-domain 3C-SiC layers with high crystalline quality[2]. The layer has an n-type doping with a nitrogen concentration of $10^{15} - 10^{16}\,\mathrm{cm}^{-3}$ and its surface is perpendicular to the $\langle 111 \rangle$ direction. The 3C-SiC layer is mechanically separated and chemical mechanical polished on both sides to leave a $730\,\mu\mathrm{m}$ thick, light-yellow-colored and freestanding layer. To isolate single defects, we irradiate a sample of this material at room temperature with 2 MeV electrons from a Cockcroft-Walton accelerator at a fluence of $5 \times 10^{12}\,\mathrm{cm}^{-2}$. We perform a subsequent anneal in Ar gas at 750° C, which causes the vacancies to migrate and form divacancies. To generate a sample with a large ensemble of defects present within our laser spot size, we use the same preparation procedures but irradiate at a larger fluence of $1 \times 10^{15}\,\mathrm{cm}^{-2}$. On the single defect sample, we lithographically pattern a Ti/Au loop antenna (10/200 nm thickness) on the surface for microwave spin rotations.

Our 4H-SiC material is grown as an on-axis epilayer on a 4H-SiC substrate using hot wall chemical vapor deposition. We mechanically separate the epilayer to form a freestanding layer, and chemical mechanical polish the surface. This process yields a $120\,\mu\mathrm{m}$ layer of single crystal 4H-SiC that has a very low unintentional dopant density of about $5 \times 10^{13}\,\mathrm{cm}^{-3}$, as determined by a mercury probe measurement. In order to create a low defect density sample where single defects can be isolated, we irradiate a sample of this material at a fluence of $5 \times 10^{12}\,\mathrm{cm}^{-2}$. We subsequently anneal both samples at 745° C in Ar gas for 30 minutes as before.

The samples are mounted on a cold finger inside a helium flow cryostat (Janis ST-500), which is incorporated into a homebuilt scanning confocal microscopy apparatus. This apparatus uses a 975 nm laser (ThorLabs, PL980P330J) for off-resonant excitation and a set of optical filters to collect the red-shifted photoluminescence above 1064 nm and record it using a two-channel superconducting nanowire detector system (SingleQuantum EOS) with approximately 80% quantum efficiency near 1300 nm. Further details of this setup are described in Ref. [3].

In our density functional theory calculations, we simulate a 512-atom 3C-SiC supercell embedding a single divacancy defect. The Brillouin zone is sampled at the Γ-point, single particle states are expressed in a 420eV cutoff plane wave basis, and atomic cores are treated with the projector-augmented-wave method [4] as implemented in VASP [5]. We use Heyd-Scuseria-Ernzerhof functionals (HSE06) hybrid functional to account for many-body exchange-correlation effects [6–11]. The atomic geometry is relaxed by using the force criterion of $0.01\,\mathrm{eV/Å}$. Contributions from

the convergent spin density as well as core state polarization are included in the hyperfine tensor calculations [12].

S2. ADDITIONAL DIVACANCY CHARACTERIZATION DATA

The data in Fig. S1 are additional characterizations of single divacancies in our low-fluence 3C-SiC and 4H-SiC samples. Fig. S1a shows the background-subtracted and spectrally resolved luminescence from a single 3C-SiC divacancy in our sample taken with using a 150 g/mm grating in a 0.3 m spectrometer. This spectrum has a zero phonon line (ZPL) at 1106 nm, which is consistent with the emission of the L3 center seen in Ref. [13]. We also find that photoluminescence excitation (PLE) measurements on single 3C-SiC divacancy defects have a center wavelength of (1106.45 ± 0.07) nm, as determined by the wavemeter, where the error reported is a 95% confidence interval. The photoluminescence spectrum shown in Fig. S4a is recorded from the high-fluence sample as a sequence of exposures with a 600g/mm grating, which are then stitched together and intensity-calibrated using a calibrated white light source (OceanOptics DH-2000-CAL). This procedure yields a photoluminescence spectrum with high spectral resolution. The ZPL seen in this spectrum is at (1106.23 ± 0.10) nm, with the reported error again as a 95% confidence interval. These ZPLs are all consistent with each other, and the slight discrepancy between the PLE and PL spectrum values in excess of the reported error is likely due to an underestimate of our systematic wavelength calibration error present in the PL measurement. The point of reporting these values is to show that the dominant defect we observe in both single defect and ensemble measurements of 3C-SiC is the L3 defect (which we identify in the main text as the neutral divacancy). We note that this ZPL is lower than the 1118 nm ZPL reported in Ref. [14], and that the origin of this discrepancy is unknown to us.

The higher resolution of the stitched spectrum can be seen in the inset of Fig. S4a, where we zoom in on the data to show the phonon sideband structure. This structure can be predicted from first-principles models, and so our intensity- and wavelength-calibrated spectrum may be useful for studying differences in the electron-phonon coupling behavior of 3C-SiC divacancies relative to the diamond NV center [15]. From a quantitative analysis of this spectrum, we find that the zero-phonon line (ZPL) contains (7.3 ± 0.3)% of the total spectrum intensity at the 95% probability level. This value, called the Debye-Waller factor, is larger than the (5.3 ± 1.1)% measured previously on a single (kk) divacancy in 4H-SiC, and the 3.2% measured on an ensemble of diamond NV centers [15]. The Debye-Waller factor has practical relevance to experiments that

use photon interference (e.g. distant entanglement schemes) because photons in the sideband must be discarded due to distinguishability. Obtaining faster success rates in defect-based entanglement schemes is an open problem, and because of the multiple binomial successes needed for a heralded entanglement [16,17], small increases in this value can produce large increases in the success rate. We also note a small peak at 1219.2 nm whose origin is not clear to us. This feature seems sharper than the wide phonon sidebands, and its shift of $(837.3 \pm 1.2)\,\mathrm{cm}^{-1}$ from the ZPL is near the 3C-SiC TO phonon mode of $796\,\mathrm{cm}^{-1}$ seen in Raman spectra measurements of 3C-SiC grown in the same way as the sample used in this work [1]. These values differ beyond our uncertainty, however, and so the origin of this peak is still unclear to us.

The background-subtracted countrates observed on single divacancies in 3C- and 4H-SiC are shown Fig. S1b. We note that these countrates are a significant improvement over the previous observations in 4H-SiC done in Ref. [3], and are the result of improvements to our apparatus and sample processing. Specifically, the photon detector in that work had a quantum efficiency of about 28%, while the detector used for most measurements in this work has an efficiency of about 80%. In addition, the 4H-SiC material we use in this work originates from the same wafer as the previous work, but the bright surface fluorescence induced by the mechanical surface polish used there has been eliminated by using a chemical mechanical polish instead. This eliminates the background surface fluorescence, and allows us to easily observe defects within about a micron or two of the surface. This was not possible in the previous work, where a $10 - 30\%$ loss in collection efficiency was estimated from having to observe defects deeper in the sample where the objective optics suffer more aberration. Our emission-rate data are taken on different inequivalent divacancies in different polytypes within the same experimental apparatus. They show that the emission rates of divacancies are comparable between polytypes, and that they are comparable to similar experiments on diamond NV centers.

The main text displays a Rabi oscillation recorded on a single 3C-SiC divacancy using off-resonant polarization and readout. We repeated this experiment at different microwave driving powers to verify that we observe the scaling of frequency predicted by Rabi's formula. The Rabi frequencies we observe as a function of microwave driving power are shown in Fig. S1c. The solid line is a fit where the expected square-root dependence is fixed and the only free parameter is an overall scaling factor. The good agreement between the solid line and the data verifies the expected dependence on microwave power. The inset shows a pulsed ODMR spectrum recorded on this defect with the expected $\approx 7.5\%$ contrast, and the solid line is a fit to two Lorentzians with a constant offset that also shows good agreement with the data. From this measurement, we extract

$D = (1.3358 \pm 0.0015)\,\mathrm{GHz} = (4.456 \pm 0.005) \times 10^{-2}\,\mathrm{cm}^{-1}$, which is near $D = 4.2 \times 10^{-2}\,\mathrm{cm}^{-1}$ reported for the L3 defect in Ref. [13].

Finally, Fig. S1d shows the two quadrature channels of a Ramsey measurement done on the same 3C-SiC divacancy. This measurement linearly adjusts the phase of our IQ modulator as a function of the wait time to give apparent oscillations in the signal. The two channels are 90° out-of-phase and are fit simultaneously as a single sinusoid in each channel with the same exponential decay. This measurement reveals an inhomogeneous spin coherence time of $T_2^* = 1.8\,\mu\mathrm{s}$, which is the longest value that we measured on single 3C-SiC divacancies. On measurements of different defects, we observed as low as $T_2^* = 200\,\mathrm{ns}$ on a particular divacancy, but generally found $T_2^* = 1 - 1.5\,\mu\mathrm{s}$ on average.

S3. HYPERFINE MEASUREMENT AND ANALYSIS DETAILS

Generally, it is easy for us to find single defects that are not strongly coupled to ^{29}Si nor ^{13}C nuclei because of the low natural abundance of Si (4.7%) and C (1.1%) nuclei with a non-zero spin. It is somewhat rare to find a divacancy with ^{13}C coupling (about a 3.4% chance, since the degeneracy in the lattice is threefold), but easier to find a divacancy with ^{29}Si coupling because of both higher isotopic abundance and higher lattice site degeneracy (about 35%, for the threefold and sixfold degeneracy lattice sites combined). In the main text, we locate two single L3/Ky5 defects that display obvious strong coupling to nearby ^{29}Si and ^{13}C nuclei in continuous wave optically detected magnetic resonance (ODMR) measurements. After locating the two defects, we take multiple ODMR measurements on each defect and use them to reconstruct their hyperfine tensors. Fig. S2 illustrates the 3C-SiC unit cell with the *ab initio* spin isosurfaces of the neutral divacancy represented by orange lobes. The $\mathrm{Si_{IIa}}$ and $\mathrm{Si_{IIb}}$ sites are explicitly labeled, and the $\mathrm{C_I}$ sites are the three nearest-neighbor carbon sites where the spin density is localized. In experiment, we obtain the best frequency resolution and the highest readout contrast by using pulsed ODMR techniques, where the spin is initialized with a fixed-length laser pulse, and a microwave pulse calibrated to produce a π rotation is then applied to the spin before reading it out with another laser pulse. Fig. S3 shows several pulsed ODMR sweeps measured on a single L3/Ky5 defect that is strongly coupled to a nearby ^{13}C nucleus. At different applied magnetic fields and angles, the six electron-nuclear states hybridize in slightly different ways, and will be slightly sensitive to different principal axes of the hyperfine tensor. The dependence of the observed transition frequencies will be strongest near the $m_s = 0$ to $m_s = -1$ anticrossing, but due to an amplifier limitation, we

use only magnetic fields and angles that produce transition frequencies above 700 MHz. In this higher frequency range, up to eight transitions can be seen, and although the electron-nuclear hybridization's dependence on the hyperfine tensor is more subtle away from the anticrossing, a global fit to multiple ODMR datasets taken on the same defect allows us to infer the hyperfine tensors reported in the main text.

Specifically, we fit Gaussian peaks to all of the pulsed ODMR datasets simultaneously (a global fit), where the peak frequencies are determined by numerical diagonalization of the total Hamiltonian. For each ODMR spectrum, the eight peak amplitudes are free parameters while the width of each peak is a single dataset-specific free parameter. These choices incorporate the fact that certain transitions can be too weak to observe, and that the width of each peak within a sweep should be the same since the π-pulse used for spin rotation is the same length at every frequency. The error on each point within a dataset is treated as uniform and set to a sweep-specific estimate of the photon shot noise. Weak normal priors are on the values of the hyperfine tensor elements for the ^{29}Si analysis (with a mean μ for $A_{xx},\ A_{yy},\ A_{zz}$ of 10 MHz and standard deviation $\sigma = 10$ MHz) to improve convergence of the fit but do not affect our final inferred values, and wide uniform priors are used for the ^{13}C analysis. The value of the crystal field splitting in the ground state (D_{gs}) is set with a strong normal prior of (1.3358 ± 0.0015) GHz, based on a separate ODMR measurement of a defect that does not display strong coupling to any nearby nuclei (Fig. S1c, inset), while the value of E is set to be less than 0.5 MHz since we have not observed any E splitting in any experiment. Uniform unbounded priors (e.g. no priors) are used for the magnetic field strength and angle, and we verified that the final fit values match our approximately known field angles and strengths.

As mentioned, the locations of the peaks are determined from numerically diagonalizing the spin Hamiltonian. Following Ref. [18], the spin Hamiltonian incorporates the electronic and nuclear Zeeman effects, the zero field splitting, and the hyperfine interaction. In this model, $A_{xx},\ A_{yy},\ A_{zz}$, and the angle θ are free parameters that are shared globally between all datasets taken on the same defect. When we fit this model to our data, the uncertainty associated with A_{yy} is quite large, which likely indicates our experiments are too far away from the anticrossing to obtain any meaningful inference of this parameter; we neglect it in our subsequent analyses by setting $A_{yy} = A_{xx}$ because their magnitudes should be similar and it reduces the dimensionality of the model.

Our interest is in the hyperfine parameters A_{zz}, A_{xx}, and θ, and ideally our estimates would be computed by integrating out every other parameter and examining the marginal probability distributions $P(A_{zz})$, $P(A_{xx})$ and $P(\theta)$. The dimensionality of our model is 110-120 parameters, so standard numerical integration techniques are infeasible and we instead use a stochastic technique

called Markov Chain Monte Carlo (MCMC) [19–21]. This technique generates samples from the full posterior probability distribution, which we use to compute the mean and credible intervals of the hyperfine tensor's parameters that are reported in the main text. Briefly, the MCMC algorithm we use works by initializing a set of points in parameter space, which are then evolved according to an adaptive proposal technique and a Metropolis-Hastings acceptance rule. This process generates a set of Markov chains whose equilibrium distribution is identical to the posterior probability distribution. This approach has the advantage of accurately propagating our experimental uncertainties through a very non-linear model, and has a tradeoff of requiring substantially larger computational effort because of the large number of iterations required for convergence. A standard single-chain MCMC algorithm cannot be easily parallelized, but our specific MCMC code uses multiple chains in parallel and a multiple-try Metropolis-Hastings acceptance rule that uses multiple proposals to update the position of a single chain. To speed up our analysis via parallelization, we implement the computation of the posterior probability density in a custom NVIDIA CUDA kernel running on a GTX 980Ti graphics processing unit [22]. At the beginning of the algorithms's execution, the ODMR data are transferred once to the global memory of the GPU. At each iteration, an array of N parameter vectors is copied to the GPU memory, and the computation of the probability density at each parameter vector is distributed to a specific computational block in the GPU. Within each block, the 6×6 Hamiltonian is computed and then diagonalized in a single GPU thread using Jacobi iteration. The eigenvalues are sorted and the eight $m_s = 0$ to $m_s = \pm 1$ transition frequencies are computed from these eigenvalues. Finally, multiple GPU threads compute and sum the normal log-likelihoods, and the function returns an array of posterior probability values to the CPU. Our GPU implementation gives a speedup between $28 - 41\times$ relative to our CPU-only implementations, and the speed of our analysis becomes limited instead by overhead from other parts of the algorithm.

We analyze the ensemble data with the same Hamiltonian model using the single data point of the observed ESEEM frequency in the periodogram of Fig. S4c. The uncertainty in the applied magnetic field and angle in the experiment are used as strong priors for the corresponding model parameters. The width of the peak in the periodogram is used as the approximate 1σ uncertainty. Since only one data point is used, the marginal distributions of A_{xx}, A_{zz}, and θ are very wide (high uncertainty), but the distribution of $A_z = (A_{zz}^2 \cos(\theta)^2 + A_{xx}^2 \sin(\theta)^2)$, the projection of the hyperfine tensor along the z-axis, is much more localized. We determine $A_z = (9.1 \pm 0.7)$ MHz and (12.4 ± 0.3) MHz for the $^{29}\mathrm{Si}_{\mathrm{IIa}}$ and $^{29}\mathrm{Si}_{\mathrm{IIb}}$ nuclei sites, respectively, from the ensemble ESEEM data, where the uncertainties are 95% probability intervals. This is in good agreement with the

$^{29}\mathrm{Si}_{\mathrm{IIa}}$ single spin data analysis where we find $A_z = (9.1 \pm 0.2)\,\mathrm{MHz}$.

S4. FLUORESCENCE LIFETIME MODELING

In this experiment, a CW pulse at 975 nm is applied to reset the spin, a microwave pulse is used to rotate the spin into a particular superposition, and finally a sub-picosecond-length pulse of 920 nm light is applied to excite the spin. The photoluminescence from this short pulse is time-resolved using a PicoHarp 300 time-correlated photon counting card. The idea behind this measurement is to create an excitation that is short enough to be comparable to an ideal delta function. That way, only states that can be optically populated by a direct transition out of the ground state are populated after the pulse, and the luminescence recorded at short durations reflects only relaxation out of these states. We assume that only one excited state is populated, since only a single zero phonon line at 1106 nm is observed under excitation with 980 nm light, and the 920 nm pulse is within the absorption band of this transition. In this setting, we assume that any multi-exponential behavior reflects different non-radiative decays out of the excited state. We perform a preliminary analysis of some initial time-resolved data by computing a nested model comparison for a single exponential model versus a biexponential model [23]. These data have the beginnings of the ghost pulses removed and had a lower signal-to-noise ratio than our final data. This means that the ghost pulses are still within the noise and thus that they can be well-described without explicit modeling of the ghost pulses that we perform in the final analysis model described below. The nested model comparison we use is a statistical test used for selecting between two competing models that compares the χ^2 statistic of each model's best fit with the respective numbers of degrees of freedom. For our data, this calculation gives $p = 1.2 \times 10^{-16}$, which indicates that the test strongly rejects the null hypothesis of the data being generated by a single exponential model versus a biexponential model. For simplicity, our final analysis uses a biexponential model and we do not test models with a higher number of exponentials.

We now describe certain systematic effects present in the raw time-correlated photon counting data and our efforts to account for them. The main difficulty with our measurement of the divacancy's fluorescence decay is that the sub-picosecond pulses used to excite the defect originate from the $\approx 76\,\mathrm{MHz}$ pulse train emitted from our Ti:Sapphire laser (Mira 900, Coherent) that is imperfectly extinguished by our electro-optic modulator. This means that the raw time-resolved photoluminescence data has small "ghost pulses" that appear about every 13 ns after the main pulse we intend to apply. The optical lifetime of the 3C-SiC divacancy is about 19 ns, so a ghost pulse

will be present in the signal even before a single $1/e$ decay time has elapsed. Multi-exponential fitting in the presence of noise can be difficult because it is mathematically ill-conditioned problem, so the presence of the ghost pulses will bias our analysis procedure if we do not explicitly account for them [24,25]. A second difficulty is a short ≈ 2 ns decay that can be observed at the beginning of the pulse. This short-lived fluorescence may originate from the optics of our confocal microscopy apparatus or background defects in the sample not seen in the spatial scans. This non-ideality is common in these types of experiments, and the simplest way to deal with it is to remove a few nanoseconds of data when the pulse begins and analyze only the later data where this background fluorescence has decayed away. This fast decay should also be present at reduced amplitude in the ghost pulses. To mitigate the short-lived fluorescence, we preprocess the data by removing six 2.5 ns long sections of data spaced 13.3 ns apart. This is the origin of the gaps in the data shown in Fig. 4b of main text and in Fig. S5.

The main part of our model is a sum of two exponential decays, where the amplitudes and lifetimes of each are free parameters, and the second decay's lifetime is constrained to be greater than the first in order to break the model's degeneracy. Six replicas of the biexponential decay are added to this main decay, and each is modulated by an amplitude scaling factor and delayed by an integer multiple of the Ti:Sapphire's repetition rate. These terms explicitly account for the ghost pulses. It is well known that EO modulators used as pulse pickers produce ghost pulses of different amplitudes, and the height of the first ghost pulse is not generally the same as the subsequent pulses. Thus, in our model, one scaling factor is used for the first ghost pulses, while a second factor is used for all subsequent pulses. We also add a constant offset to account for background photoluminescence. The posterior probability of this model is computed simultaneously (a global fit) over multiple datasets that prepare the spin into different spin superpositions in the ground state. The lifetimes, laser repetition rate, and ghost pulse amplitude scaling parameters are shared between datasets, since they should not change between different runs, while the decay amplitudes and constant offsets are free parameters for each dataset. Finally, we fit the model to the data using a MCMC approach similar to that described in the discussion of the hyperfine data analysis. We use a normally distributed error whose magnitude at each point is the square root of the total bin counts multiplied by a scaling factor η. This factor allows for the model to account for error beyond the Poisson-distributed counting error, and can be a goodness-of-fit measure. After fitting, we find this parameter is less than 1.05. A value near 1 indicates that our model does a good job of capturing the actual random error we observe in experiment. We use only weak non-informative priors on all parameters. The parameters and uncertainties quoted in the main text are the mean

and 95% credible intervals of the respective marginal probability distributions, which are computed using the output of the MCMC sampler.

The data and model fits are shown as points and solid lines in the upper panel of Fig. S5. The data are shown out to longer delays here compared to the main text. This makes the difference between the two decay curves harder to see by eye, but allows us to illustrate that the residuals in the lower panel are approximately normally distributed across the entirety of the decays. The residual plot is another measure of the goodness-of-fit of our model. It visually shows that the model does a good job of capturing the main decay and ghost pulses and that our per-point experimental uncertainties are appropriate.

S5. RATE EQUATION MODELING

We use a set of five levels with first order transition rates between them to understand the 3C-SiC divacancy's simplified spin and orbital dynamics. This model makes the assumption that the excited state is only a single triplet manifold and that the $m_s = \pm 1$ levels in both the ground and excited states are each single levels. Any spin singlet levels that participate in the dynamics are modeled as a single fifth level. In Appendix A of the main text, we use two experiments to constrain the rates between levels. The best-fit rates, shown in Table S1, are determined by a numerical optimization of the rate-equation model output to the normalized photoluminescence output at eleven different laser excitation powers [26,27]. This optimization is constrained by the two lifetimes and ground-state spin polarization inferred from the biexponential decay experiment.

S6. SPIN RESONANCE CLASSIFICATION ALGORITHM

There are four inequivalent forms of divacancy in the 4H-SiC lattice that have their own similar but distinct optical and magnetic resonance properties. Single divacancies in 4H-SiC appear as bright spots in a confocal photoluminescence spatial scan, as in Fig. S6, and their form can be identified by a characteristic spin resonance signature seen in an ODMR measurement. Our experiment that infers the excited state Hamiltonian parameters from PLE measurements of (hh) and (kk) divacancies requires finding defects with a wide range of transverse strain values because different Hamiltonian parameters are more significant in different ranges of transverse strain. Because the strain experienced by a given defect is random, this can require a search through a large number of candidate bright spots in order to locate defects that are both the correct form of divacancy,

determined via ODMR, and experience a particular amount of strain, via a PLE measurement that is conditional on the form of defect. Having to manually identify the defect from its spin resonance measurement means a search over many defects can be tedious and time consuming.

We automate this search by recording candidate bright spots into a list of coordinate tuples and entering a loop where an ODMR measurement is taken on a defect, a custom classification algorithm determines its type based on the ODMR data, and finally the resonant laser is adjusted to perform a PLE measurement whose particular experimental settings depends on the form of defect determined in the classification step. The general strategy of automating measurements that occur in a particular order is not new, but we describe our classification algorithm here since it may be a useful component in future efforts where automated classification with high accuracy is required, e.g. the search for a particular form of divacancy with the lowest strain, highest T_2^* or T_2, etc.

Neural networks have made recent strides in, among other things, recognizing complicated objects in digital images. Convolutional neural networks are particularly successful in these tasks because their local connectivity makes them sensitive to features that arise in a small patch of pixels located near each other without the large dimensionality of a densely connected layer. We train a convolutional neural network to classify 80-pixel ODMR measurements that we take on bright spots in our 4H-SiC sample as an (hh) divacancy, a (kk) divacancy, a (kh) divacancy, or "noise". We do not observe any (hk) divacancies in our experiments, similar to Ref. [3], so we do not include them as a classification case. Examples of these four cases are shown in Fig. S7a. We create a neural network in the Python package Keras with a Theano backend that uses a common design where convolution and pooling layer pairs are repeated, ReLU activation functions are used, and two densely connected layers form the final part of the network [28,29]. Fig. S7b shows our particular design that when trained on 150 ODMR traces achieves an estimated classification accuracy of 98.0%, as determined by a 10-fold cross validation analysis. Twenty of the training examples are "noise" datasets that we artificially generate (a form of data augmentation), and this roughly balances the number of each class seen by the network in training. Our choice of small number of total layers and a densely connected layer of only 20 nodes reflects the fact that these images are collected under homogeneous conditions (similar measurement times and the same applied magnetic field) and inherently have a very simple structure of peaks or dips that unambiguously determine the type defect. This algorithm represents a way to achieve fast classification of simple spin resonance data with high accuracy and does not require the creation of handcrafted features. We found in practice that the simpler technique of a least squares fit of two peaks to the data and

identifying the defect based on the sign of the amplitudes required a good initial starting point and would sometimes erroneously fit a peak to noise. Starting the fit from many initial points (> 300) improved this approach's robustness, but increased the runtime substantially, and its accuracy was still only about 90%. Areas of future exploration might be the performance of a similar neural network algorithm on data collected under different magnetic field conditions versus the amount of downsampling/maxpooling employed (which affects translation invariance), or the application of similar techniques to spin resonance data taken on other defect systems.

S7. ADDITIONAL PLE DATA AND ANALYSIS DETAILS

The excited state model we use to analyze our data has been used for nitrogen vacancy centers in diamond, and is developed from group theory arguments in two previous works [30,31]. The notation used in each work is slightly different, but the Hamiltonians are identical up to a basis/coordinate system change. We use the sum of the spin-orbit, spin-spin, and strain interaction matrices to model the excited state level behavior we observe in our PLE measurements. The notation we adopt corresponds to $\lambda_z = \lambda_z$, $\Delta = D_{\text{es}}/3$, $\Delta' = \Delta_1/2$, and $\Delta'' = \Delta_2$ in the work of Maze, et al. and $\lambda_\parallel = \lambda_z$, $D_{2,A_1} = D_{\text{es}}/3$, $D_{2,E,1} = \Delta_1/2$, and $D_{2,E,2} = \Delta_2/\sqrt{2}$ in the work of Doherty, et al.

Our inferences of λ_z, D_{es}, Δ_1 and Δ_2 originate from a Bayesian fit of the parameterized Hamiltonian's eigenvalues to the relative excited state energies seen in our PLE measurements. The relative excited state energies can be computed based on the transition resonance locations observed in experiment by considering the change in energy from a level in the ground state to a particular level in the excited state. Our experiment uses small magnetic fields of a few Gauss in order to break the degeneracy of the $m_s = \pm 1$ levels in the ground state. This splitting is only a few MHz, which is well within our observed linewidths, so we approximate the $m_s = \pm 1$ levels as degenerate and separated from $m_s = 0$ by exactly D_{gs} when preprocessing the PLE transition frequencies into relative energies. Our PLE measurement uses off-resonant excitation to polarize the ground state spin primarily into the $m_s = 0$ state, so by taking the same PLE measurement with or without microwaves applied, we can label the states as originating from $m_s = 0$ or $m_s = \pm 1$ in the ground state. We fit a sum of Lorentzians to the PLE spectra to extract the transition frequencies with statistical uncertainty estimates, and finally compute the relative energies in the excited state by subtracting D_{gs} from the transitions originating from the $m_s = \pm 1$ ground state sublevels. When pre-processing the PLE data taken on the upper branch of the excited state, the

splitting between two levels of different spin character is sometimes near D_{gs}. This means that one of the transitions of $m_s = \pm 1$ character is at the same frequency as the $m_s = 0$ transition within less than the $\approx 100\,\mathrm{MHz}$ linewidth, and it is difficult to resolve them clearly. In this case, we set both transitions to a fit of a single Lorentzian to the degenerate transition peak. This likely introduces some systematic error whose magnitude is near one linewidth, and we account for it in the conservative formulation of our likelihood function described below.

Our posterior probability function incorporates several potential sources of experimental uncertainty. The specified repeatability of our wavemeter is about 8 MHz, and similar uncertainties are quoted from the Lorentzian fitting done in the preprocessing step. We use a conservative good data/bad data likelihood that treats each point as originating from a mixture of two normal distributions. We assume the first has a standard deviation of $\sigma = 20\,\mathrm{MHz}$ (the 'good data') and the second has $\sigma = 300\,\mathrm{MHz}$ (the 'bad data'), and that the relative fraction of 'bad data' is about 5%. This approach assigns a lower weight to outlier points in a natural way, and means that rather than eliminating data with slightly worse apparent errors, the complete set is used for statistical inference [32,33]. We also treat a second possible source of systematic error that might occur when the upper and lower branches are measured in separate PLE measurements. Because they are separated in time by at least a few minutes, when repeated laser tracking and other measurements might occur, it is possible that the local charge distribution may shift the frequencies somewhat – a form of slow spectral diffusion. We performed a separate experiment on a single (kk) divacancy designed to observe spectral diffusion by repeatedly scanning the laser across a transition (see section S8 for a detailed discussion of this experiment and its analysis). No obvious jumps are seen within a $\approx 100\,\mathrm{MHz}$ linewidth, the resonance never disappeared in any scan, and our quantitative analysis sets an upper bound of about $\sigma = 18\,\mathrm{MHz}$ on the standard deviation of the jump size from scan to scan. We account for the possibility of a jump between scans by allowing for a random uniform shift between the upper and lower branch energies, which is an approximation for a small shift in transverse strain. We conservatively assign this effect a normal prior with zero mean and $\sigma = 500\,\mathrm{MHz}$.

Having formulated the probability model, we apply an MCMC algorithm to sample from the full posterior probability distribution [34]. The mean and 95% credible intervals of the marginal distributions of parameters for both (hh) and (kk) divacancies are reported in the main text. The residuals also appear to match our aforementioned likelihoods – most of the points have tight agreement within $\approx 20\,\mathrm{MHz}$ with the maximum a posteriori (MAP) estimate for both the (hh) and (kk) data, but a few outliers in the (hh) data are $100 - 250\,\mathrm{MHz}$ away. As mentioned, these points

are automatically weighted lower by our choice of likelihood. Fig. S8 shows the data and model fit for both (hh) and (kk) divacancies along with the residuals. The data plotted are shifted to cancel out the axial strain and any systematic jump between branches in accordance with the probability model's maximum a posteriori value. The probability model incorporates the weak anisotropy of transverse strain, but the solid fit lines in the upper panels fix the strain purely along a single direction for plotting purposes. The residuals in the lower panels are between the data and the Hamiltonian eigenvalues where this anisotropy is accounted for, so the residual data are slightly smaller than the mismatch between the corresponding points and solid lines in the upper panels.

To validate our model's implementation, we digitize the excited-state energy level data taken on single diamond NV center defects by Batalov et al. (Ref. [35]). Using a similar analysis, we determine $\lambda_z = (5.30 \pm 0.06)\,\mathrm{GHz}$, $D_{\mathrm{es}} = (1.42 \pm 0.04)\,\mathrm{GHz}$, $\Delta_1 = (1.58 \pm 0.02)\,\mathrm{GHz}$, and $\Delta_2 = (0.21 \pm 0.08)\,\mathrm{GHz}$ where the errors are 95% credible intervals. These estimates are in close quantitative agreement with the values of $\lambda_z = 5.3\,\mathrm{GHz}$, $D_{\mathrm{es}} = 1.42\,\mathrm{GHz}$, $\Delta_1 = 1.55\,\mathrm{GHz}$, and $\Delta_2 = 0.2\,\mathrm{GHz}$ reported by Batalov et al., which suggests our implementation and parameter definitions are correct. No explicit uncertainties were reported in that work to which we might compare our values. However, considering that the residuals relative to the model appear to have a standard deviation of about 70 MHz, and the data cover regions of low, intermediate, and high transverse strain, our uncertainties of $20 - 80\,\mathrm{MHz}$ on each parameter seem reasonable.

A. Remarks on spin mixing in the ^3E excited state manifold

A key motivation of this work is in examining the quality of the divacancy's cycling transitions for high-fidelity preparation and readout protocols [36–38]. In these protocols, specific transitions between the ground and excited states are driven repeatedly to extract as many photons as possible until a spin flip occurs and the photon emission ceases. It is essential to understand spin-mixing in the excited state since this is the primary origin of the spin flips that degrade the achievable initialization and readout fidelities, and some further technical remarks on this topic are warranted due to its importance. Our discussion here follows closely some of the results and discussion in Ref [36]. To begin, if a highly pure initialization into the $m_s = 0$ sublevel of the ground state is desired, an off-resonant laser pulse can be applied first, which polarizes the spin to $\approx 94\%$ using the divacancy's intrinsic optical cycle. Then, to eliminate the residual population in the $m_s = \pm 1$ sublevels, the transition from the $m_s = \pm 1$ ground state sublevels to the A_1 level of the excited state can be driven. This transition has a mostly $m_s = \pm 1$ character, but a small amount of

mixing gives the transition a slight $m_s = 0$ character. After repeated excitation, the defect will be shelved into the $m_s = 0$ (dark) state from either spontaneous emission or the defect's intrinsic spin-selective intersystem crossing. In Ref. [36], a preparation error of around 0.4% is quoted when using this approach. On the other hand, for readout, the E_x is predominantly of $m_s = 0$ character, so driving it when the spin is prepared as $m_s = 0$ will cause the defect to excite and spontaneously decay back to the $m_s = 0$ level in the ground state with very high probability. It can then be re-excited to extract additional photons, up to a limit determined by the amount of $m_s = \pm 1$ spin that E_x contains. Increased mixing increases the probability p of a spin flip occurring when the system spontaneously decays, and thus reduces the number of photons that can be extracted according to a geometric distribution whose expectation value is $\frac{1-p}{p}$. This is $\approx 1/p$, for small p, and if p is very small, like the measurements of Refs. [35] and [36] indicate, and as our measurements in 4H-SiC divacancies now indicate, the number of total photons that can be extracted before a spin flip is highly sensitive to this value. Thus, in single-shot detection experiments, where discrimination between $m_s = 0$ and $m_s = \pm 1$ can be done by integrating the total photon counts recorded during a single finite readout period and deciding which state was detected based on whether multiple photons or no photons were detected, the magnitude of the spin flip probability is highly important. Two effects play a main role whether detection with high-fidelity within a single measurement shot is possible: the total number of photons that are emitted when the spin is projected into $m_s = 0$ (which depends on p), and how many of those photons are actually detected. In Ref. [36], with a solid immersion lens and other improvements designed to maximize the photon detection efficiency, the average number of photons recorded for the NV center in the $m_s = 0$ was only about 6.4 photons, which meant that there was a roughly 13% chance of detecting no photons and incorrectly assigning the detected state as $m_s = \pm 1$. Because of the low collection efficiencies in our current experiment, even if spin-mixing in the excited state of the divacancy is low enough to yield a number of emitted photons near that of the diamond NV, so few of these photons will be detected that we are very far away from the single-shot detection regime. Nevertheless, if spin-mixing in the excited state is low enough, it means that with comparable detection efficiencies, the same or higher fidelities than diamond NV's should be achievable for SiC divacancies. We see no obvious reasons at the moment that the collection efficiency enhancements in Ref. [36] could not also be engineered in SiC, and so any quantitative predictions we can make about the number of photons emitted are very relevant for the prospects of single-shot detection of divacancy spins.

In the upper plot of Fig. S9, we use the (hh) divacancy excited state parameters inferred

in the main text to compute $m_s = 0$ content (the squared matrix element) of the E_x and E_y levels as a function of transverse strain. These values are the per-excitation probabilities of the spin remaining in $m_s = 0$, and their complementary values are the probabilities of a spin flip to $m_s = \pm 1$ after spontaneous emission. The E_x level is almost purely $m_s = 0$ across the entire range of strain, while the E_y level experiences an anticrossing around 9 GHz (c.f. Ref. [36]). The location of the anticrossing changes with the strain direction, while the behavior of E_x is mostly isotropic, so we mainly discuss E_x here for simplicity. As mentioned earlier, under constant driving and independent spin-flip probabilities, the total number of photons extracted before a spin flip occurs follows a geometric distribution with expectation $\langle N_{\mathrm{ph}} \rangle \approx 1/p$. This probability is small in both NV centers and SiC divacancies, so small differences can give large changes in the expected number of photons, $\langle N_{\mathrm{ph}} \rangle$. For comparison, the lower plot in Fig. S9 shows $\langle N_{\mathrm{ph}} \rangle$ as a function of transverse strain for the E_x transition in 4H-SiC (hh) and (kk) divacancies and the diamond NV center. The posterior distribution of our analysis of the excited-state data extracted from Ref. [35] is used used to create the predictions and uncertainties for the NV center. The value of p and hence $\langle N_{\mathrm{ph}} \rangle$ is very sensitive to the value of Δ_2, the term in the Hamiltonian that mixes E_x/E_y with E_1/E_2, and these predicted values explicitly illustrate the role of a smaller Δ_2 in SiC producing a higher expected number of photons. These predictions use the full posterior distribution for the SiC and diamond Hamiltonian parameters we inferred above to predict the probability mode (solid lines) and 95% prediction intervals (shaded bands) of $\langle N_{\mathrm{ph}} \rangle$. As mentioned, the number of photons that can be extracted is a key factor in high-fidelity discrimination between spin states, so this computation shows that the mixing behavior of divacancies is likely as good as or better than NV centers, where single shot readout has already been demonstrated. In practice, other mechanisms like orbital hopping between E_x and E_y or weak intersystem crossing coupling of E_x or E_y to the singlet states may reduce $\langle N_{\mathrm{ph}} \rangle$, but these important effects have not yet been characterized in SiC divacancies [39,40].

S8. REPEATED MEASUREMENTS OF PLE ON A SINGLE 4H-SiC (kk) DIVACANCY

Spectral diffusion and ionization are potentially significant issues when resonantly exciting the optical transitions in solid-state defects. A slow spectral diffusion will manifest as discrete jumps in the location of a resonance over time, while ionization manifests the disappearance of a resonance after it is excited. Fig. S10 shows the raw photon counts at each applied laser frequency in 62 sequential PLE scans performed over about a 30 minute period on a single (kk) divacancy. In these

measurements, off-resonant pumping is not applied during a scan, and is only applied about every 90 seconds between scans for a brief feedback measurement to compensate for drift in the cryostat over time. No ionization nor discrete jumps are seen in these scans, and we have never observed any jumps or ionization in any of our other measurements of divacancies. In this particular set of measurements in Fig. S10, we tune the piezo of our resonant laser to change the laser frequency by small increments and measure the photoluminescence during a 200 ms readout integration time in each scan. Immediately before the data in Fig. S10 were measured, we execute a single laser frequency scan where at each step of the piezo voltage we measure the laser frequency using our wavemeter. We fit a second order polynomial to the frequency versus piezo voltage data from this scan to calibrate this relationship, and the frequency we record for subsequent scans is determined by this fit from the voltage set point only. It is important to note that any drift in this calibration will confound the true frequency of the laser and the inferred magnitude of spectral diffusion, and hence our analysis can only produce an upper bound on the true magnitude of spectral diffusion.

In our data, no discrete spectral jumps, which would indicate a slow spectral diffusion magnitude larger than our optical linewidth, are obvious by eye. However, smaller jumps could be present within the ≈118 MHz linewidth of the optical transition. When the jump magnitude is smaller than the linewidth, its measurement is complicated by the fact that each scan is also subject to statistical noise. Individual fits of a Lorentzian to the recorded scan data will produce slightly different fit values for the center frequency, even if no actual jumps are occurring (this is essentially the meaning of a "confidence interval"). Thus, when analyzing the data using individual fits, it is unclear what portion of the observed change in the center frequency's fit value from scan to scan is due to spectral jumps and which portion is due to statistical noise alone.

To separate the effect of statistical noise from the underlying magnitude of spectral jumps, we use a Bayesian hierarchical model to analyze our experiment. In it, the Poisson-distributed photon shot noise and the spectral jumps, which we assume for simplicity follow a Gaussian distribution, are modeled separately. Explicitly, the model encodes the relationships,

$$y_{ij}|A_i,\nu_i,C_i,\Gamma,x_{ij} \sim \text{Poisson}\left(f(A_i,\nu_i,C_i,\Gamma,x_{ij})\right), \tag{1}$$

$$f(A,\nu,C,\Gamma,x) = A\frac{\left(\frac{\Gamma}{2}\right)^2}{\left(\frac{\Gamma}{2}\right)^2+(x-\nu)^2}, \tag{2}$$

$$\nu_i \sim \text{Normal}\left(\nu_{\text{grand}},\sigma_{\text{sd}}\right), \tag{3}$$

$$\nu_{\text{grand}} \sim \text{Normal}\left(0, 200\,\text{MHz}\right), \tag{4}$$

$$\sigma_{\mathrm{sd}} \sim \text{Exponential}\left(1/(120\,\mathrm{MHz})\right). \tag{5}$$

To give a plain explanation of this model, we now discuss these equations in turn. In Eq. 1, each data point, y_{ij}, where i denotes the scan number and j denotes the point number within the i^{th} scan, originates from a Poisson distribution whose intensity parameter (in units of photons) is dependent on the laser frequency, x_{ij}, and given by a function f. The function f in Eq. 2 is a Lorentzian with an amplitude A, a center frequency ν, a constant offset C, and a full-width-half-maximum Γ. In the model, A, C, and ν are allowed to vary from scan to scan, while Γ is treated as the same in every scan. As mentioned above, we assume the center frequencies, which could jump due to slow spectral diffusion, follow a Gaussian (Normal) distribution (Eq. 3). This distribution has parameters ν_{grand} and σ_{sd}, which are the grand mean of the central frequencies, and the standard deviation of central frequencies, respectively. The value of σ_{sd} represents the magnitude of the possible scan-to-scan spectral jumps in our data, and hence it is our primary interest in this analysis. Since we preprocess the data to roughly subtract the central frequency, we set the hyperparameters in the prior distribution for ν_{grand} in Eq. 4 to a mean of 0 MHz and standard deviation of 200 MHz. This simply means that we expect the average location of the resonances to be near 0 MHz, and we use a wide uncertainty of 200 MHz to keep this prior non-informative. By eye and based on our previous experience of observing no obvious jumps on other defects, the spectral jump magnitude is likely less than one linewidth ($\approx$100 MHz). To reflect this, we set the prior probability density of σ_{sd} as an exponential distribution with a mean of 120 MHz in Eq. 5. This is again much larger than our expectations so that this prior density is kept non-informative. The prior probability densities for A and C are simply uniform bounded below by 0 and above by large constants.

We sample from the posterior probability density using MCMC [34], and these samples allow us to compute an upper bound to the spectral diffusion magnitude $\sigma_{\mathrm{sd}} < 18$ MHz at the 95% probability level. This is much smaller than one linewidth, and makes it clear how this kind of analysis can segment out even small jump magnitudes from noisy data. Although no discrete jumps are seen in Fig. S10, a small drift in the resonance over the course of the 62 scans is slightly apparent. We think this frequency drift may be a slight drift in the piezo inside the laser cavity used to adjust the laser frequency. Since any piezo drift will also be included in the magnitude of σ_{sd}, this suggests the true magnitude of slow spectral diffusion is even smaller than the upper bound our model determines.

S9. PHOTOLUMINESCENCE EXCITATION OF A 4H-SiC (*kh*) DIVACANCY

We also perform photoluminescence excitation (PLE) and optically detected magnetic resonance (ODMR) measurements on a single (*kh*) divacancy in our 4H-SiC sample. This form of divacancy is oriented along the basal plane and not the c-axis of the 4H-SiC crystal. It has a reduced C_{1h} symmetry, which means that the model we adopted from C_{3v}-symmetric NV centers in diamond does not describe its excited state structure. A detailed understanding of this structure is still an open problem and outside the scope of this work, but we now describe our empirical observations and comparisons to three prior works.

In the work of Koehl, et al., ensembles of this defect were associated in optically detected magnetic resonance measurements with a bright photoluminescence line near 1.1497 eV [41]. A later work by Falk, et al. associated ensembles of this defect with two lines – one near 1.148 eV and another near 1.118 eV – and also observed the splitting of each branch into at least three sublevels with an applied strain [42]. The PLE and coherent population trapping experiments of Zwier, et al. were performed on ensembles of this defect near 1.1495 eV, and the lower energy line was not reported [43]. We measure a photoluminescence excitation signal from both energy branches on a single (*kh*) divacancy in Fig. S11. We find the central frequencies of each branch of this particular defect are at 270,695 GHz (1.1195 eV) and 278,053GHz (1.1499 eV). The data in Fig. S11 are shifted relative to these frequencies. In these measurements, the intensity of the lower branch appears to be much less ($\approx 8\times$) than the upper branch. We did not investigate the relative intensity of each line as a function of polarization, but our observation that the lower line is weaker is consistent with the $\approx 3\times$ reduction seen by Falk, et al. and the at least 5x reduction between the 1.1195 eV and 1.1497 eV lines seen by Koehl, et al.

We measure the ODMR spectrum of this single (*kh*) divacancy at low field in Fig. S12a, and observe a resonance at 1.311 GHz and another at 1.358 GHz. The difference in ODMR peak heights/widths is due to different coupling strengths of each resonance to a fixed-strength microwave field: each transition experiences a different amount of saturation and power broadening. To calibrate this effect out, we perform separate Rabi oscillation measurements to determine the microwave strengths needed to drive Rabi oscillations at a frequency of about 4.04 MHz on the two respective resonances. When we perform a PLE measurement of the lower branch while applying these calibrated microwaves to the ground state spin transition at $f = 1.311$ GHz, we observe an approximately $(23 \pm 5)\%$ increase in the strength of the PLE line (shown as the upper trace in the left panel of the figure) relative to the case when no microwaves are applied (not shown),

where the error reported is a 95% confidence interval. When we drive microwaves at a frequency of $f = 1.358\,\mathrm{GHz}$ and the driving strength is adjusted according to our previous calibration, we observe no change in the signal relative to the no microwave case (not shown). This behavior is unlike the case in c-axis oriented divacancies where ground state spin populations in either $m_s = +1$ or $m_s = -1$ levels will transition to the A_1, A_2, E_1, and E_2 excited state levels in roughly equal proportion because of the high degree of mixing between $m_s = \pm 1$ in those levels. To verify the (kh) divacancy PLE strength's dependence on the $f = 1.358\,\mathrm{GHz}$ microwave driving frequency, we apply microwaves again with the same strength but at $f = 1.45\,\mathrm{GHz}$ (shown in the figure), which is far from any magnetic resonance transition. We observe no change the signal strength relative to the no-microwave case, which implies that the change we see when driving at $f = 1.358\,\mathrm{GHz}$ is real and not an artifact of any microwave heating of the sample.

Our measurement of PLE on upper branch displays a more distinct dependence on microwave driving. When no microwaves are applied, we observe a single resonance at 278,053 GHz. When we apply calibrated-strength microwaves ($f_{\mathrm{Rabi}} = 4.04\,\mathrm{MHz}$) to the $f = 1.311\,\mathrm{GHz}$ transition, the strength of this PLE resonance increases by $(18 \pm 7)\%$ and a new PLE resonance appears about 1 GHz lower in energy. When we apply calibrated-strength microwaves to the $f = 1.358\,\mathrm{GHz}$ transition, the strength of the main PLE resonance decreases by about $(13 \pm 6)\%$, shifts lower in frequency by about 80 MHz, and the same PLE resonance 1 GHz lower in energy appears again with equal strength to the previous PLE measurement. The small shift downward is visible by eye in the plot and is unlikely to be a result of heating because the microwave power used at $f = 1.358\,\mathrm{GHz}$ is less than the power used at $f = 1.311\,\mathrm{GHz}$ where no shift was observed.

In both cases, the strength of the line seen with no microwave driving increases in intensity when microwaves at $f = 1.311\,\mathrm{GHz}$ are applied, and decreases in intensity when microwaves at $f = 1.358\,\mathrm{GHz}$ are applied. This is consistent with the peak and dip at 1.311 and 1.358 GHz, respectively, in the ODMR spectrum collected using the off-resonant laser. This difference in the direction of the intensity change is unlike the c-axis oriented divacancies, where both transitions produce an intensity change in the same direction. We do not understand the origins of the differences in ODMR behavior seen between defects, and more detailed studies of the specific branching ratios out of the excited state will likely shed light on this question.

As a final demonstration, we apply a resonant laser to the upper branch transition that is visible when no microwaves are applied and monitor its fluorescence in a Rabi oscillation experiment (Fig. S12b). This is the same experimental pulse sequence as the Rabi oscillations recorded in Fig. 3b of the main text on a c-axis oriented divacancy. Rabi oscillations are clearly visible and have a

contrast of about 44%. This is substantially lower than the 94% contrast observed in the main text experiment, but is likely from another transition with different spin character being located within the PLE resonance. This experiment serves as a proof-of-principle that resonant readout can also work in the basal plane (kh) divacancy, and we leave a more careful exploration for future studies.

References

[1] V. Jokubavicius, G. R. Yazdi, R. Liljedahl, I. G. Ivanov, J. Sun, X. Liu, P. Schuh, M. Wilhelm, P. Wellmann, R. Yakimova, and M. Syväjärvi, "Single Domain 3C-SiC Growth on Off-Oriented 4H-SiC Substrates," Cryst. Growth Des. **15**, 2940 (2015).

[2] V. Jokubavicius, G. R. Yazdi, R. Liljedahl, I. G. Ivanov, R. Yakimova, and M. Syväjärvi, "Lateral Enlargement Growth Mechanism of 3C-SiC on Off-Oriented 4H-SiC Substrates," Cryst. Growth Des. **14**, 6514 (2014).

[3] D. J. Christle, A. L. Falk, P. Andrich, P. V Klimov, J. Ul Hassan, N. T. Son, E. Janzén, T. Ohshima, and D. D. Awschalom, "Isolated Electron Spins in Silicon Carbide with Millisecond-Coherence Times," Nat. Mater. **14**, 160 (2015).

[4] P. E. Blöchl, "Projector Augmented-Wave Method," Phys. Rev. B **50**, 17953 (1994).

[5] G. Kresse, "From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method," Phys. Rev. B **59**, 1758 (1999).

[6] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," Phys. Rev. Lett. **77**, 3865 (1996).

[7] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Erratum: Hybrid Functionals Based on a Screened Coulomb Potential (Journal of Chemical Physics (2003) 118 (8207))," J. Chem. Phys. **124**, 219906 (2006).

[8] J. Heyd and G. E. Scuseria, "Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd-Scuseria-Ernzerhof Screened Coulomb Hybrid Functional," J. Chem. Phys. **121**, 1187 (2004).

[9] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid Functionals Based on a Screened Coulomb Potential," J. Chem. Phys. **118**, 8207 (2003).

[10] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals," J. Chem. Phys. **125**, 224106 (2006).

[11] K. Szász, V. Ivády, E. Janzén, and A. Gali, "First Principles Investigation of Divacancy in SiC Polytypes for Solid State Qubit Application," Mater. Sci. Forum 778–780, **499** (2014).

[12] K. Szasz, T. Hornos, M. Marsman, and A. Gali, "Hyperfine Coupling of Point Defects in Semiconductors by Hybrid Density Functional Calculations: The Role of Core Spin Polarization," Phys. Rev. B **88**, 75202 (2013).

[13] N. T. Son, E. Sörman, W. M. Chen, C. Hallin, O. Kordina, B. Monemar, E. Janzén, and J. L. Lindström, "Optically Detected Magnetic Resonance Studies of Defects in Electron-Irradiated 3C SiC Layers," Phys. Rev. B **55**, 2863 (1997).

[14] G. Calusine, A. Politi, and D. D. Awschalom, "Silicon Carbide Photonic Crystal Cavities with Integrated Color Centers," Appl. Phys. Lett. **105**, 11123 (2014).

[15] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van De Walle, "First-Principles Theory of the Luminescence Lineshape for the Triplet Transition in Diamond NV Centres," New J. Phys. **16**, 73026 (2014).

[16] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson, "Heralded Entanglement Between Solid-State Qubits Separated by Three Metres," Nature **497**, 86 (2013).

[17] S. D. Barrett and P. Kok, "Efficient High-Fidelity Quantum Computation Using Matter Qubits and Linear Optics," Phys. Rev. A **71**, 60310 (2005).

[18] V. Ivády, K. Szász, A. L. Falk, P. V. Klimov, D. J. Christle, E. Janzén, I. A. Abrikosov, D. D. Awschalom, and A. Gali, "Theoretical Model of Dynamic Spin Polarization of Nuclei Coupled to Paramagnetic Point Defects in Diamond and Silicon Carbide," Phys. Rev. B **92**, 115206 (2015).

[19] C. Geyer, "Practical Markov Chain Monte Carlo," Stat. Sci. 7, 473 (1992).

[20] J. S. Liu, F. Liang, and W. H. Wong, "The Multiple-Try Method and Local Optimization in Metropolis Sampling," J. Am. Stat. Assoc. **95**, 121 (2000).

[21] E. Laloy and J. A. Vrugt, "High-Dimensional Posterior Exploration of Hydrologic Models Using Multiple-Try DREAM(ZS) and High-Performance Computing," Water Resour. Res. **48**, 1 (2012).

[22] J. Nickolls, I. Buck, M. Garland, and K. Skadron, "Scalable Parallel Programming with CUDA," AMC Queue **6**, 40 (2008).

[23] J. H. Steiger, A. Shapiro, and M. W. Browne, "On the Multivariate Asymptotic Distribution of Sequential Chi-Square Statistics," Psychometrika 50, 253 (1985).

[24] A. A. Istratov and O. F. Vyvenko, "Exponential Analysis in Physical Phenomena," Rev. Sci. Instrum. **70**, 1233 (1999).

[25] J. R. Lakowicz, "Principles of Fluorescence Spectroscopy" (2006).

[26] K. Schittkowski, "NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm with Distributed and Non-Monotone Line Search" (2012).

[27] R. H. Byrd, J. Nocedal, and R. A. Waltz, "Knitro: An Integrated Package for Nonlinear Optimization," in Large-Scale Nonlinear Optim. (2006), pp. 35–59.

[28] F. Chollet, "Keras," GitHub Repos. https://github.com/fchollet/keras (2015).

[29] The Theano Development Team, R. Al-Rfou, G. Alain, A. Almahairi, C. Angermueller, D. Bahdanau, N. Ballas, F. Bastien, J. Bayer, A. Belikov, A. Belopolsky, Y. Bengio, A. Bergeron, J. Bergstra, V. Bisson, J. B. Snyder, N. Bouchard, N. Boulanger-Lewandowski, X. Bouthillier, A. de Brébisson, O. Breuleux, P.-L. Carrier, K. Cho, J. Chorowski, P. Christiano, T. Cooijmans, M.-A. Côté, M. Côté, A. Courville, Y. N. Dauphin, O. Delalleau, J. Demouth, G. Desjardins, S. Dieleman, L. Dinh, M. Ducoffe, V. Dumoulin, S. E. Kahou, D. Erhan, Z. Fan, O. Firat, M. Germain, X. Glorot, I. Goodfellow, M. Graham, C. Gulcehre, P. Hamel, I. Harlouchet, J.-P. Heng, B. Hidasi, S. Honari, A. Jain, S. Jean, K. Jia, M. Korobov, V. Kulkarni, A. Lamb, P. Lamblin, E. Larsen, C. Laurent, S. Lee, S. Lefrancois, S. Lemieux, N. Léonard, Z. Lin, J. A. Livezey, C. Lorenz, J. Lowin, Q. Ma, P.-A. Manzagol, O. Mastropietro, R. T. McGibbon, R. Memisevic, B. van Merriënboer, V. Michalski, M. Mirza, A. Orlandi, C. Pal, R. Pascanu, M. Pezeshki, C. Raffel, D. Renshaw, M. Rocklin, A. Romero, M. Roth, P. Sadowski, J. Salvatier, F. Savard, J. Schlüter, J. Schulman, G. Schwartz, I. V. Serban, D. Serdyuk, S. Shabanian, É. Simon, S. Spieckermann, S. R. Subramanyam, J. Sygnowski, J. Tanguay, G. van Tulder, J. Turian, S. Urban, P. Vincent, F. Visin, H. de Vries, D. Warde-Farley, D. J. Webb, M. Willson, K. Xu, L. Xue, L. Yao, S. Zhang, and Y. Zhang, "Theano: A Python Framework for Fast Computation of Mathematical Expressions," arXiv: 1605.02688, 19 (2016).

[30] J. R. Maze, A. Gali, E. Togan, Y. Chu, A. Trifonov, E. Kaxiras, and M. D. Lukin, "Properties of Nitrogen-Vacancy Centers in Diamond: The Group Theoretic Approach," New J. Phys. **13**, 25025 (2011).

[31] M. W. Doherty, N. B. Manson, P. Delaney, and L. C. L. Hollenberg, "The Negatively Charged Nitrogen-Vacancy Centre in Diamond: The Electronic Solution," New J. Phys. **13**, 25019 (2011).

[32] G. E. P. Box and G. C. Tiao, "A Bayesian Approach to Some Outlier Problems," Biometrika **55**, 119 (1968).

[33] D. Sivia and J. Skilling, "Data Analysis: A Bayesian Tutorial" (1998).

[34] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC

Hammer," Publ. Astron. Soc. Pacific **125**, 306 (2013).

[35] A. Batalov, V. Jacques, F. Kaiser, P. Siyushev, P. Neumann, L. J. Rogers, R. L. McMurtrie, N. B. Manson, F. Jelezko, and J. Wrachtrup, "Low Temperature Studies of the Excited-State Structure of Negatively Charged Nitrogen-Vacancy Color Centers in Diamond," Phys. Rev. Lett. **102**, 195506 (2009).

[36] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. Alkemade, and R. Hanson, "High-Fidelity Projective Read-Out of a Solid-State Spin Quantum Register," Nature **477**, 574 (2011).

[37] P. Tamarat, N. B. Manson, J. P. Harrison, R. L. McMurtrie, A. Nizovtsev, C. Santori, R. G. Beausoleil, P. Neumann, T. Gaebel, F. Jelezko, P. Hemmer, and J. Wrachtrup, "Spin-Flip and Spin-Conserving Optical Transitions of the Nitrogen-Vacancy Centre in Diamond," New J. Phys. **10**, 45004 (2008).

[38] N. B. Manson, J. P. Harrison, and M. J. Sellars, "Nitrogen-Vacancy Center in Diamond: Model of the Electronic Structure and Associated Dynamics," Phys. Rev. B **74**, 104303 (2006).

[39] M. L. Goldman, A. Sipahigil, M. W. Doherty, N. Y. Yao, S. D. Bennett, M. Markham, D. J. Twitchen, N. B. Manson, A. Kubanek, and M. D. Lukin, "Phonon-Induced Population Dynamics and Intersystem Crossing in Nitrogen-Vacancy Centers," Phys. Rev. Lett. **114**, 145502 (2015).

[40] L. J. Rogers, R. L. McMurtrie, M. J. Sellars, and N. B. Manson, "Time-Averaging Within the Excited State of the Nitrogen-Vacancy Centre in Diamond," New J. Phys. **11**, 63007 (2009).

[41] W. F. Koehl, B. B. Buckley, F. J. Heremans, G. Calusine, and D. D. Awschalom, "Room Temperature Coherent Control of Defect Spin Qubits in Silicon Carbide," Nature **479**, 84 (2011).

[42] A. L. Falk, P. V. Klimov, B. B. Buckley, V. Ivády, I. A. Abrikosov, G. Calusine, W. F. Koehl, Á. Gali, and D. D. Awschalom, "Electrically and Mechanically Tunable Electron Spins in Silicon Carbide Color Centers," Phys. Rev. Lett. **112**, 187601 (2014).

[43] O. V Zwier, D. O'Shea, A. R. Onur, and C. H. van der Wal, "All-Optical Coherent Population Trapping with Defect Spin Ensembles in Silicon Carbide," Sci. Rep. **5**, 10931 (2015).

Parameter	Rate (MHz)
k_{31}	43
k_{32}	2
k_{41}	5
k_{42}	43
k_{35}	11
k_{45}	21
k_{51}	50
k_{52}	0
$P_1/(P_1 + P_2)$	0.95

TABLE S1: Transition rates determined by a fit of the rate-equation model to the time-resolved photoluminescence data.

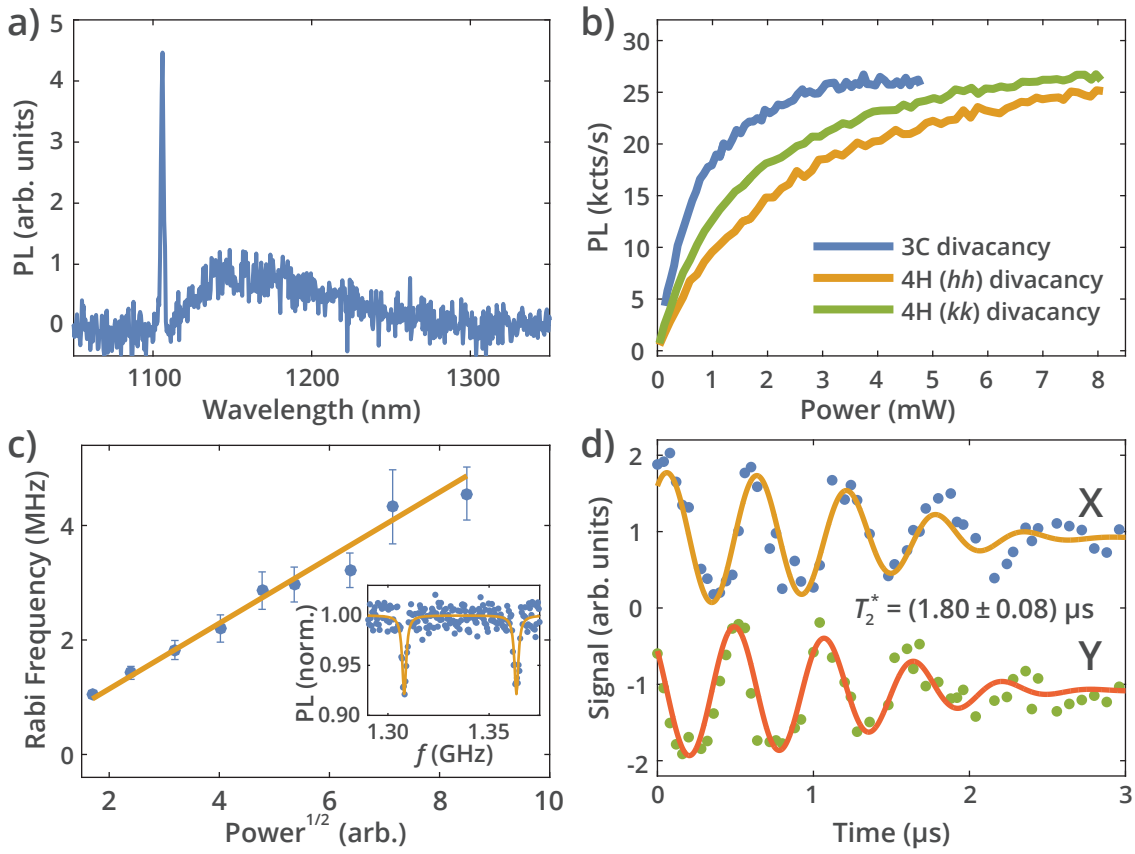

FIG. S1: Photoluminescence and microwave characterizations of single divacancies in 3C-SiC and 4H-SiC. (a) Photoluminescence spectrum recorded on a single 3C-SiC divacancy. (b) Photoluminescence intensity in kcts/s as a function of laser power incident on the back of the microscope objective for the 3C-SiC divacancy and the (hh) and (kk) forms of divacancy in 4H-SiC. The background luminescence recorded near each defect for each intensity has been subtracted. The countrate measured on each divacancy form appears to saturate near 27 kcts/s. (c) Frequency of Rabi oscillations observed on a single 3C-SiC divacancy as a function of the square root of the applied microwave power. Inset: Pulsed optically detected magnetic resonance spectrum for the same divacancy showing the spin sublevel transitions split by a magnetic field of a few Gauss. (d) Differential photoluminescence showing the X and Y quadrature readout channels of a Ramsey measurement taken on the same divacancy as in panel c. The channels are fit simultaneously to sinusoids that are out-of-phase by 90° multiplied by a decay function $\exp(-t/T_2^*)$, and we extract that $T_2^* = (1.80 \pm 0.08)\,\mu$s. The phase of each readout channel is rotated at a rate of 2 MHz to induce the oscillations shown.

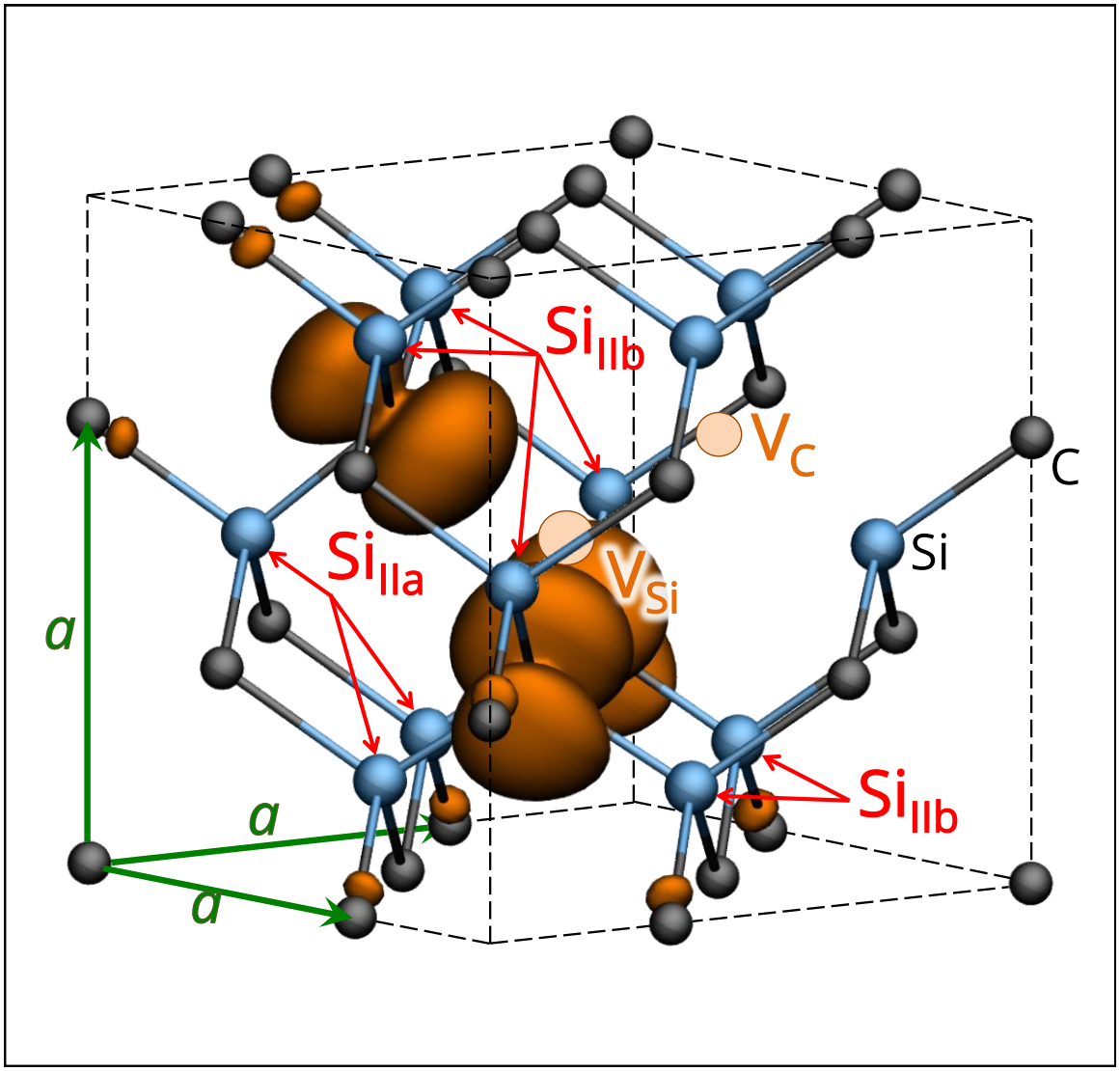

FIG. S2: Ball and stick diagram of the 3C-SiC unit cell with a divacancy defect. The missing atoms that comprise the divacancy are shown as V_{Si} and V_C. The calculated spin density is represented by the orange-lobe isosurfaces and is primarily localized at the dangling bonds of the C atoms nearest to the Si vacancy. The three Si_{IIa} and six Si_{IIb} sites are indicated by red arrows. The C_I sites are the three first-neighbour carbon sites where the spin density is localized.

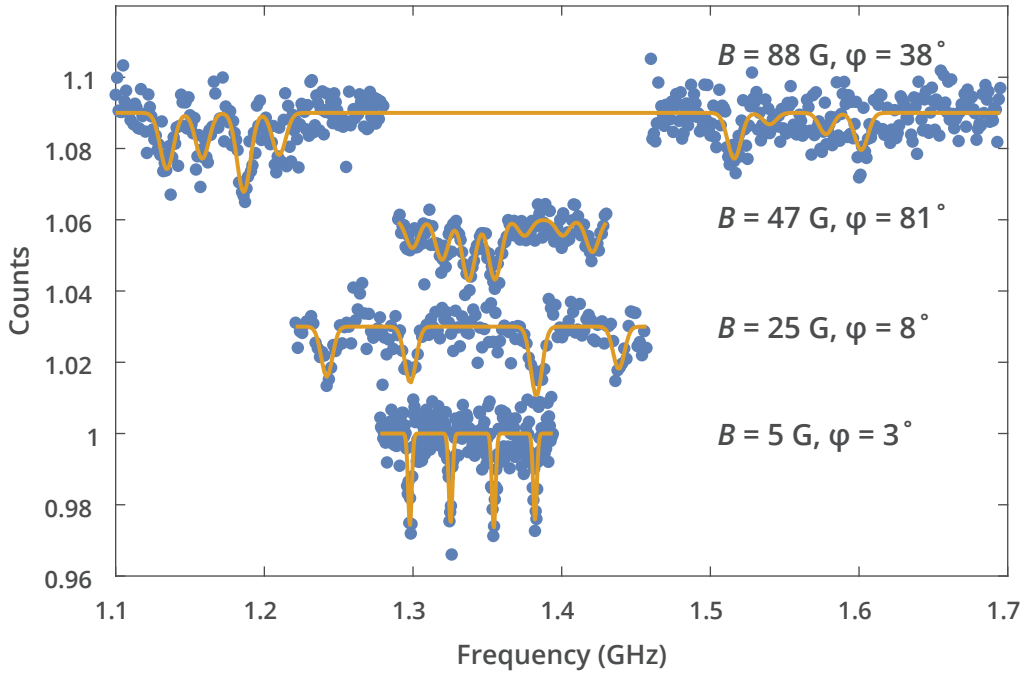

FIG. S3: Low-field pulsed ODMR spectra recorded on a single 3C-SiC L3/Ky5 defect strongly coupled to a nearby ^{13}C nucleus. Each sweep is recorded at different a magnetic field strength (B) and angle relative to the defects symmetry axis (φ). The sweeps are normalized to the photoluminescence emitted when the applied microwaves are not resonant with any transition, and offset by 0.03 to allow for comparison. The solid lines are the best fit curves output from a global fit to these and other pulsed ODMR sweeps.

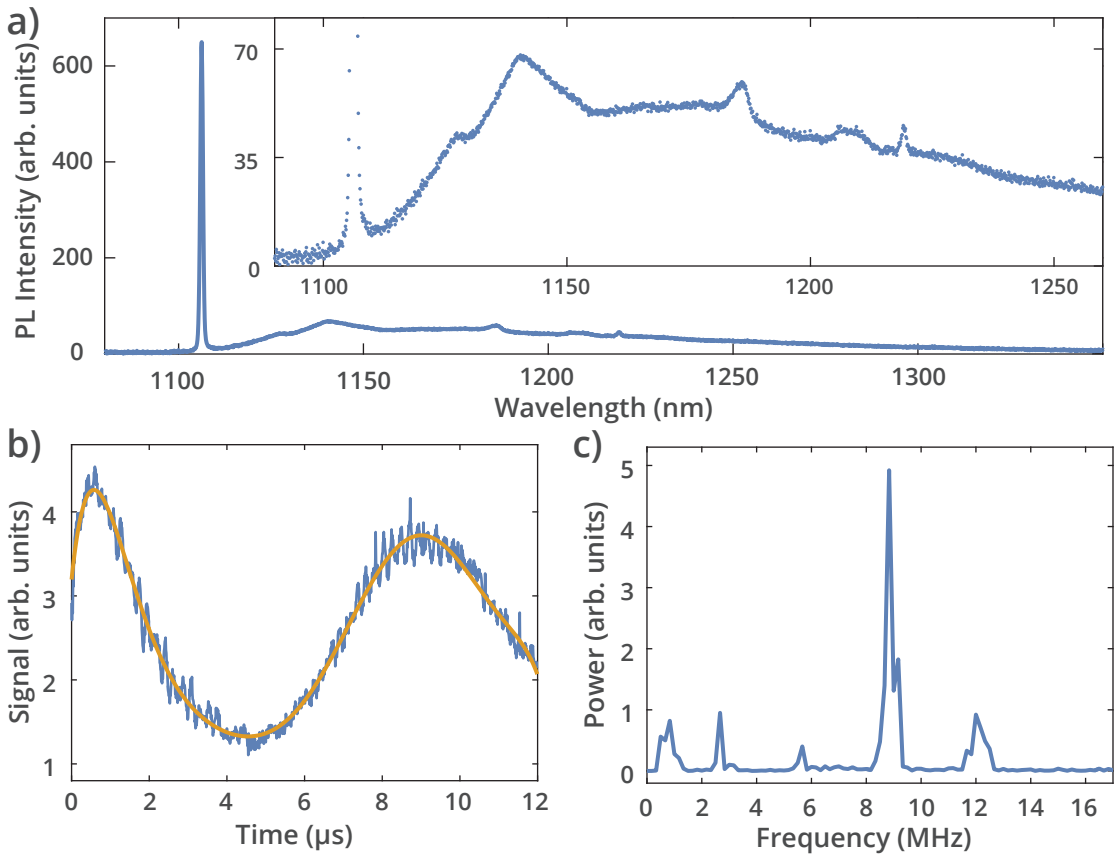

FIG. S4: Photoluminescence spectrum and ESEEM data recorded from the 3C-SiC sample electron-irradiated with a $1 \times 10^{15}\,\mathrm{cm}^{-2}$ fluence to create an ensemble of divacancies. (a) Calibrated high-resolution photoluminescence spectrum taken on the $1 \times 10^{15}\,\mathrm{cm}^{-2}$ fluence 3C-SiC ensemble sample. The neutral divacancy at 1106 nm appears to be the dominant defect present since it is the only obvious zero phonon line we observed. Inset: Zoomed in view to highlight the phonon sideband structure of the divacancy's emission. (b) Short-duration Hahn echo measurement taken on an ensemble of 3C-SiC divacancies at $B = 253\,\mathrm{G}$ at $T = 20\,\mathrm{K}$. The slow oscillations are the Larmor precession of the ^{29}Si and ^{13}C nuclei, while the faster oscillations are the hyperfine couplings of these nuclei to the divacancy's electronic spin. The orange curve is a fit of an 8^{th} order polynomial that is used to capture the slow oscillations for subtraction. (c) The power spectral density of the fast oscillations that remain after subtraction of the Larmor precession. The spectral peaks near 9 MHz and 12.5 MHz were fit with Lorentzians for estimating the value of A_z in Table I of the main text for the two ^{29}Si sites.

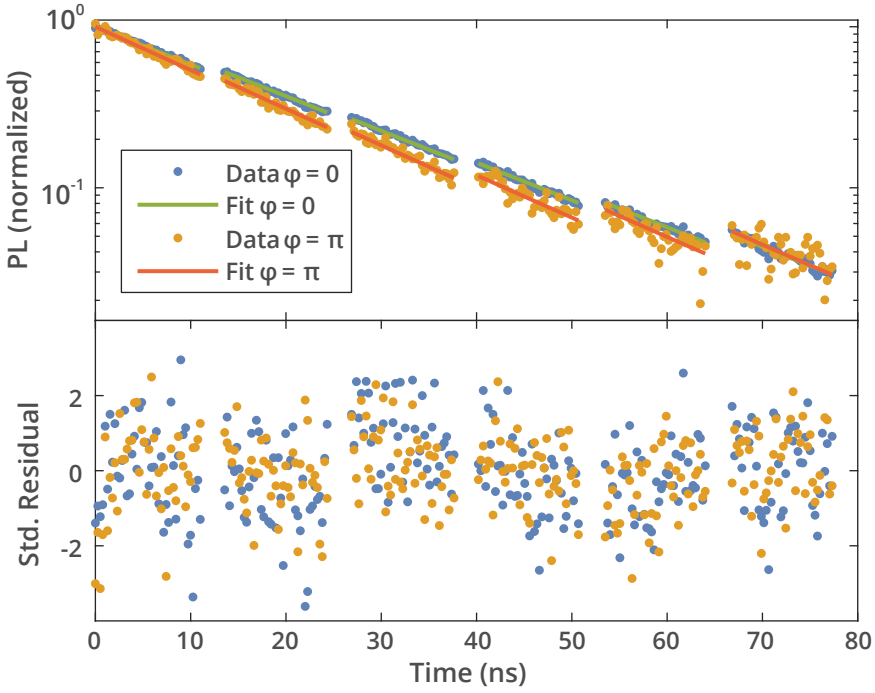

FIG. S5: Time-resolved photoluminescence of a single 3C-SiC divacancy. (upper panel) Normalized photoluminescence decay of a single 3C-SiC divacancy prepared in the $m_s = 0$ state and with a π-pulse applied to the $m_s = -1$ transition. The solid lines are the biexponential decay model photoluminescence at the best fit value. (lower panel) Residuals of the same data and model divided by the square root of the raw bin counts in order to standardize them for comparison. These values follow an approximate normal distribution with unit variance, as expected.

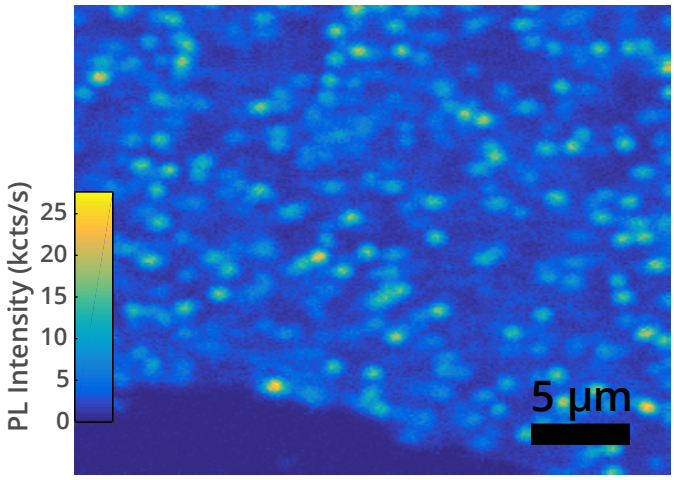

FIG. S6: Two-dimensional photoluminescence spatial scan of 4H-SiC sample. This sample is irradiated with 2 MeV electrons to a fluence of $5 \times 10^{12}\,\mathrm{cm}^{-2}$, as described in the main text. With such a low fluence, single divacancy defects are easily visible as isolated bright spots. The dark area of the image in the lower left is the edge of a lithographically patterned metal loop that we use for microwave spin control.

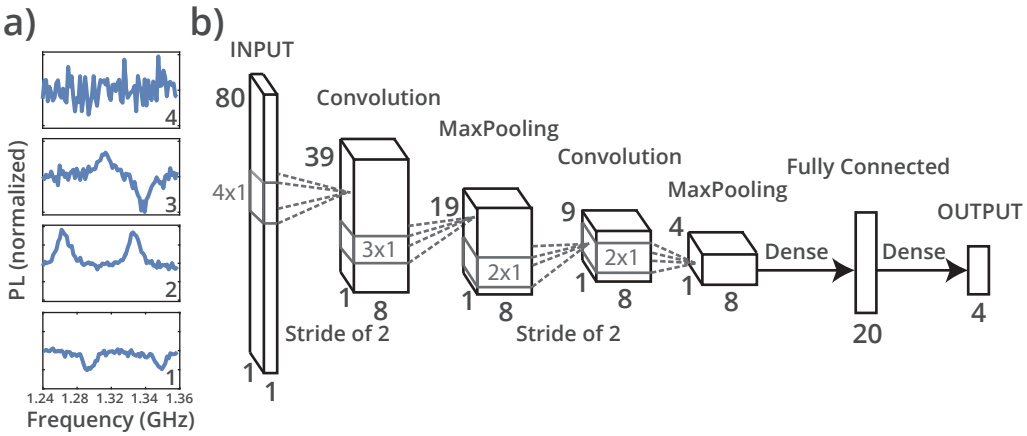

FIG. S7: Convolutional neural network for spin resonance classification. (a) Normalized ODMR measurements representative of the four cases used in the classification algorithm: 1–(hh) divacancy, 2–(kk) divacancy, 3–(kh) divacancy, and 4–Noise/indeterminate. (b) Diagram showing each layer of the feedforward neural network used for classification. Numbers exterior to the layer boxes indicate the shape of the neuron layout, while the grey interior numbers indicate the sizes of the convolutional or downsampling filters applied between layers.

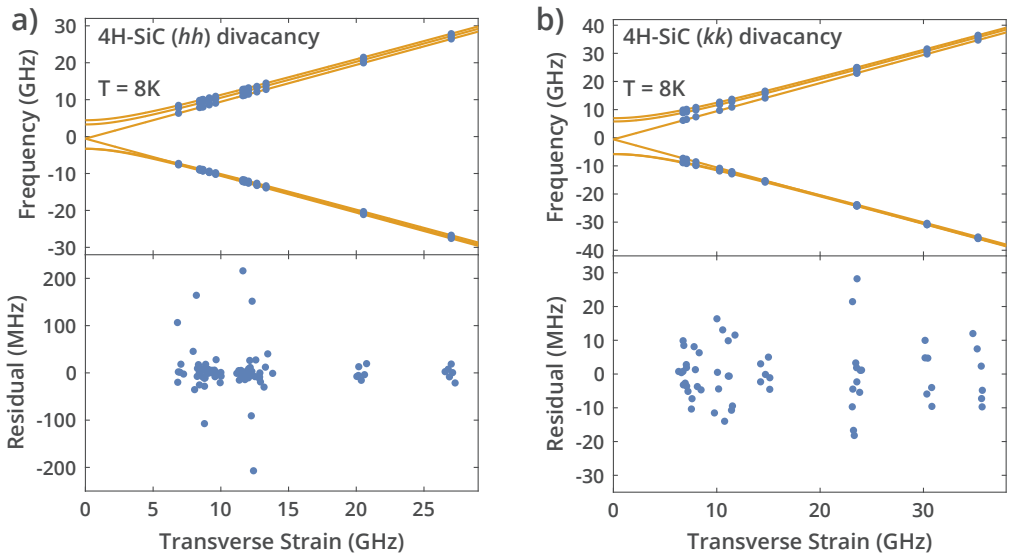

FIG. S8: 4H-SiC divacancy excited state energies versus transverse strain. a) Photoluminescence excitation resonances recorded on separate (hh) divacancies at $T = 8\,\mathrm{K}$. b) Photoluminescence excitation resonances recorded on separate (kk) divacancies at $T = 8\,\mathrm{K}$. The solid lines in the upper panel of both subfigures are the Hamiltonian eigenvalues at the maximum a posteriori (MAP) value for transverse strain purely along the x direction. The residuals are computed using the eigenvalues at the MAP where both the magnitude and direction of the transverse strain of each defect are taken into account. The value of transverse strain for each datum in the residual plots is jittered by 1 GHz to reduce overplotting.

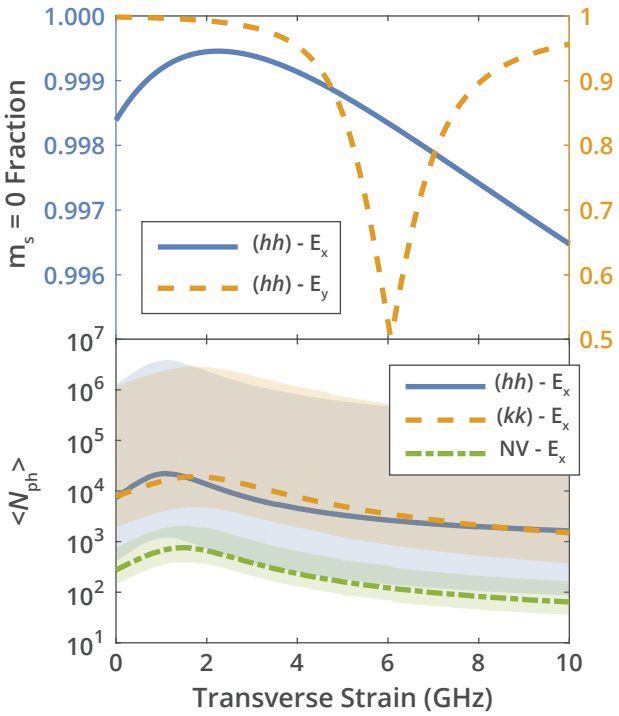

FIG. S9: Predicted excited-state spin-mixing behavior for 4H-SiC divacancies and the diamond NV center versus transverse strain. (upper panel) The squared probability amplitude of $m_s = 0$ for the E_x and E_y excited states of (hh) divacancies. For this form of divacancy, the E_y level experiences an anticrossing near 6 GHz. The location of this anticrossing depends on the direction of transverse strain, and the direction here is chosen to be consistent with Ref. [36]. The behavior of the E_x state is less sensitive to this direction, so we study only E_x for different defects in the lower panel. (lower panel) The expected total number of photons emitted from single defects when exciting the $m_s = 0$ to E_x optical transition as a function of transverse strain. This quantity is the expectation value of a geometrically distributed number of emitted photons, $1/(1 - p_{\mathrm{m}_s=0})$, where $p_{\mathrm{m}_s=0}$ the $m_s = 0$ squared probability amplitude. The direction of transverse strain for each defect is the same as the upper panel. Solid lines are the mode of the posterior predictive distribution for $1/(1 - p_{\mathrm{m}_s=0})$, while the shaded areas indicate the 95% prediction intervals.

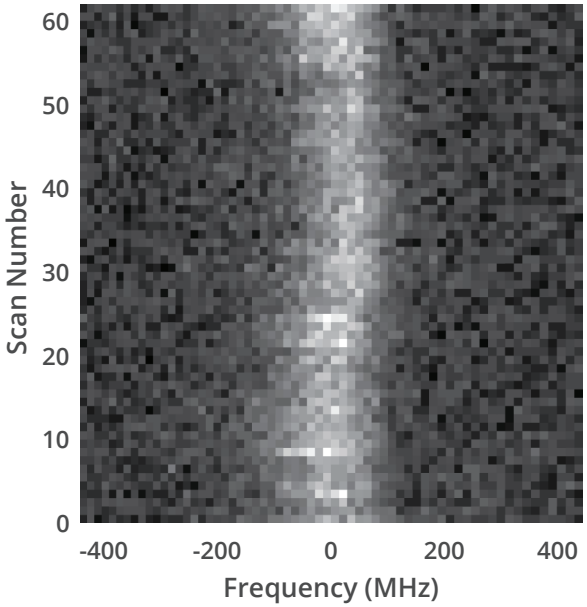

FIG. S10: Photoluminescence excitation scans on a single 4H-SiC (kk) divacancy over time as a function of the applied resonant laser frequency.

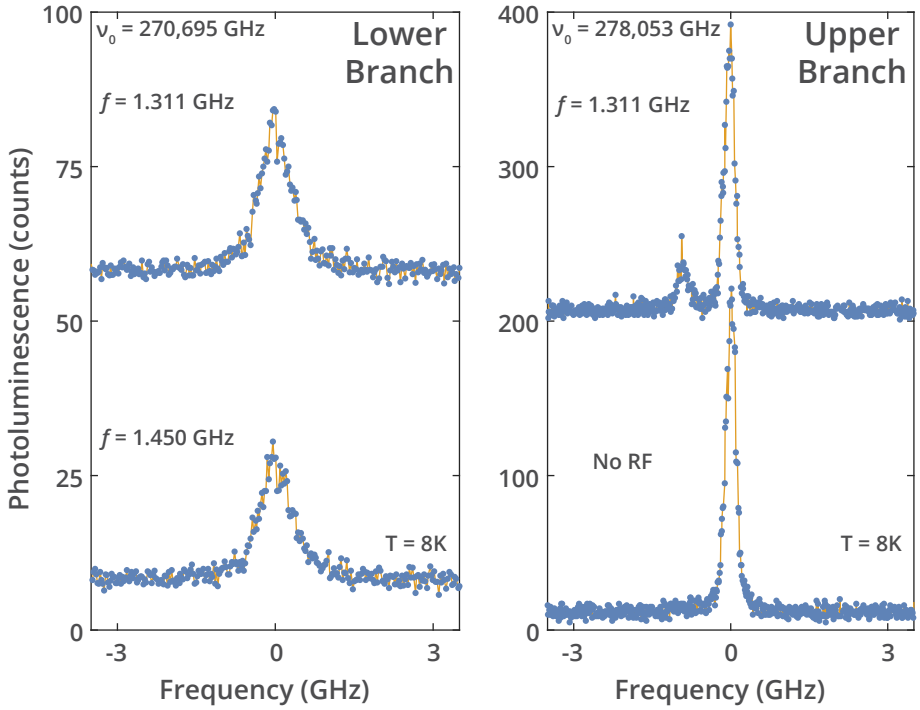

FIG. S11: Photoluminescence excitation measurements taken on a single (kh) divacancy in 4H-SiC. (left panel) The lower trace shows the counts measured when sweeping the resonant laser across the lower branch (relative to $\nu_0 = 270{,}695\,\mathrm{GHz}$) at a laser power of about 314 nW when applying microwaves at a non-resonant frequency of 1.45 GHz. The upper trace is the taken under the same conditions but with the microwave driving frequency at 1.311 GHz, which is resonant with the lower frequency microwave transition in the ground state. This peak is about $(23.5 \pm 5.4)\%$ higher than the lower trace. (right panel) The lower trace shows the counts measured when sweeping the resonant laser across the upper branch (relative to $\nu_0 = 278{,}053\,\mathrm{GHz}$) at a laser power of about 165 nW when no microwaves are applied. The upper trace is taken under the same conditions but with the microwave driving frequency again at 1.311 GHz.

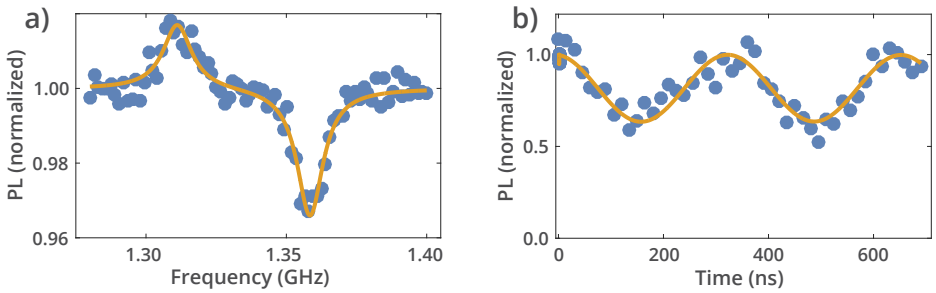

FIG. S12: Magnetic resonance signals recorded on a single (kh) divacancy in 4H-SiC. a) Continuous-wave optically detected magnetic resonance taken using the off-resonant laser. The trace shows the two expected peaks at 1.311 GHz and 1.358 GHz. b) Resonant readout of Rabi oscillations detected by monitoring the fluorescence of the transition near 278,053 GHz that is visible when no microwaves are applied. The readout contrast of the oscillations is approximately 45%.

---

* Corresponding author.
Email address: awsch@uchicago.edu