Probing Spin Dynamics on Diamond Surfaces Using a Single Quantum Sensor

Dwyer, Bo L.; Rodgers, Lila V.H.; Urbach, Elana K.; Bluvstein, Dolev; Sangtawesin, Sorawis; Zhou, Hengyun; Nassab, Yahia; Fitzpatrick, Mattias; Dobrovitski, V. V.; More Authors

DOI
10.1103/PRXQuantum.3.040328

Publication date
2022

Document Version
Final published version

Published in
PRX Quantum

Featured in Physics

Probing Spin Dynamics on Diamond Surfaces Using a Single Quantum Sensor

Bo L. Dwyer,[1,†] Lila V.H. Rodgers,[2,†] Elana K. Urbach,[1,†] Dolev Bluvstein,[1] Sorawis Sangtawesin,[3] Hengyun Zhou,[1] Yahia Nassab,[2,‡] Mattias Fitzpatrick,[2] Zhiyang Yuan,[2] Kristiaan De Greve,[1,§] Eric L. Peterson,[1] Helena Knowles,[1,¶] Tamara Sumarac,[1] Jyh-Pin Chou,[4] Adam Gali,[5,6] V.V. Dobrovitski,[7] Mikhail D. Lukin,[1] and Nathalie P. de Leon[2,*]

[1] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

[2] *Department of Electrical and Computer Engineering, Princeton University, Princeton, New Jersey 08544, USA*

[3] *School of Physics and Center of Excellence in Advanced Functional Materials, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand*

[4] *Department of Physics, National Changhua University of Education, Changhua 50007, Taiwan*

[5] *Institute for Solid State Physics and Optics, Wigner Research Centre for Physics, POB 49, Budapest H-1525, Hungary*

[6] *Department of Atomic Physics, Institute of Physics, Budapest University of Technology and Economics, Műegyetem rakpart 3., Budapest 1111, Hungary*

[7] *QuTech and Kavli Institute of Nanoscience, Delft University of Technology, Delft 2628 CD, Netherlands*

(Received 24 March 2022; revised 27 June 2022; accepted 6 October 2022; published 14 December 2022)

Understanding the dynamics of a quantum bit's environment is essential for the realization of practical systems for quantum information processing and metrology. We use single nitrogen-vacancy (NV) centers in diamond to study the dynamics of a disordered spin ensemble at the diamond surface. Specifically, we reduce the density of "dark" surface spins to interrogate their contribution to the decoherence of shallow NV center spin qubits. When the average surface spin spacing exceeds the NV center depth, we find that the surface spin contribution to the NV center free induction decay can be described by a stretched exponential with variable power n. We show that these observations are consistent with a model in which the spatial positions of the surface spins are fixed for each measurement, but some of them reconfigure between measurements. In particular, we observe a depth-dependent critical time associated with a dynamical transition from Gaussian ($n = 2$) decay to $n = 2/3$, and show that this transition arises from the competition between the small decay contributions of many distant spins and strong coupling to a few proximal spins at the surface. These observations demonstrate the potential of a local sensor for understanding complex systems and elucidate pathways for improving and controlling spin qubits at the surface.

DOI: 10.1103/PRXQuantum.3.040328

I. INTRODUCTION

Understanding and controlling the environment of quantum bits is a central challenge in solid-state quantum science and engineering. Nitrogen-vacancy (NV) color centers in diamond have emerged as a promising platform for numerous applications in quantum sensing and information processing [1–3]. While bulk NV centers constitute spin qubits with exceptional coherence properties, even under ambient conditions, several applications require shallow NV centers to be placed within nanometers of the diamond surface. In particular, these shallow NV centers can be used for nanoscale sensing applications, enabling detection of individual electron spins [4,5], molecules [6], or nuclei [7]. However, shallow NV centers exhibit increased decoherence due to surface defects [8–10], including ubiquitous "dark" unpaired electron spins on the diamond surface [4,5]. Prior studies have demonstrated that these dark spins themselves can have long spin relaxation and coherence times, making them potentially useful as "quantum reporters" in sensing applications [5,7]. While these surface spins have been observed and studied for several years [4,5,7,10–12], their origin and properties

*Correspondence and requests for materials should be addressed to npdeleon@princeton.edu

†These authors contributed equally to this work.

‡Current address: Department of Chemical Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

§Current address: Inter-university Microelectronics Centre, Imec, Kapeldreef 75, Leuven, Belgium.

¶Current address: Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0HE, United Kingdom

are not well understood, limiting the potential applications of shallow NV centers to harness them as a resource for sensing and simulation.

Motivated by these considerations, in this work, we systematically investigate the quantum dynamics of over 100 individual shallow NV centers proximal to the diamond surface, using surface treatments to adjust the density of surface spins, σ. We find that the surface spin contribution to the coherence decay, measured across many NV centers in numerous samples, provides insight into the nature of these defects. Specifically, we are able to probe the dilute ($d_{\mathrm{NV}} < 1/(2\sqrt{\sigma})$) surface spin limit, and we observe that the coupling between a NV center and the surface spin bath increases monotonically with proximity to the surface rather than exhibiting increased variance. We also observe that very few ($< 2\%$) of NV centers exhibit coherent coupling to single surface spins. Both observations are at odds with the prevailing physical model for surface spins, wherein their positions are fixed and only their spin orientation can change. Moreover, this large dataset permits us to observe a dynamical crossover from Gaussian ($n = 2$) decay to stretched exponential decay ($n < 1$) that is indicative of a dilute two-dimensional (2D) spin bath, averaged over many spatial configurations. We show that the crossover time is a generic feature of disordered ensembles, which has broad relevance for other solid-state qubit platforms.

II. DEER DECAY AND SURFACE SPIN DENSITY

Our experiments are depicted in Fig. 1(a). NV centers are produced by ion implantation [6,10] at depths $d_{\mathrm{NV}} = 3$–20 nm from the surface, under a number of different surface conditions (Appendix A). We use proton NMR measurements to accurately determine the implanted NV depth [13]. The surface-spin-NV interaction is studied using double electron-electron resonance (DEER) on a large, random selection of individual NV centers in each sample [14]. The DEER pulse sequence, illustrated in Fig. 1(b), is similar to that of a spin echo, except that microwave π pulses are applied to both the NV center spin and the surface spins midway through the free evolution time [5]. This allows for selective interrogation of the flipped spins, while the signal due to nonflipped spins is canceled to first order by virtue of the echo. These experiments are performed at a magnetic field of approximately 300 Gauss so that the NV center and surface spin transitions are frequency resolved, and with surface spin Rabi frequencies > 60 MHz so that π pulses are short compared to both the surface spin relaxation and dephasing (Appendix D).

Figure 1(c) shows an example DEER dataset. In order to remove effects that limit NV center spin coherence in the absence of surface spins, we normalize the DEER decay signal by the coherence signal obtained through a spin echo

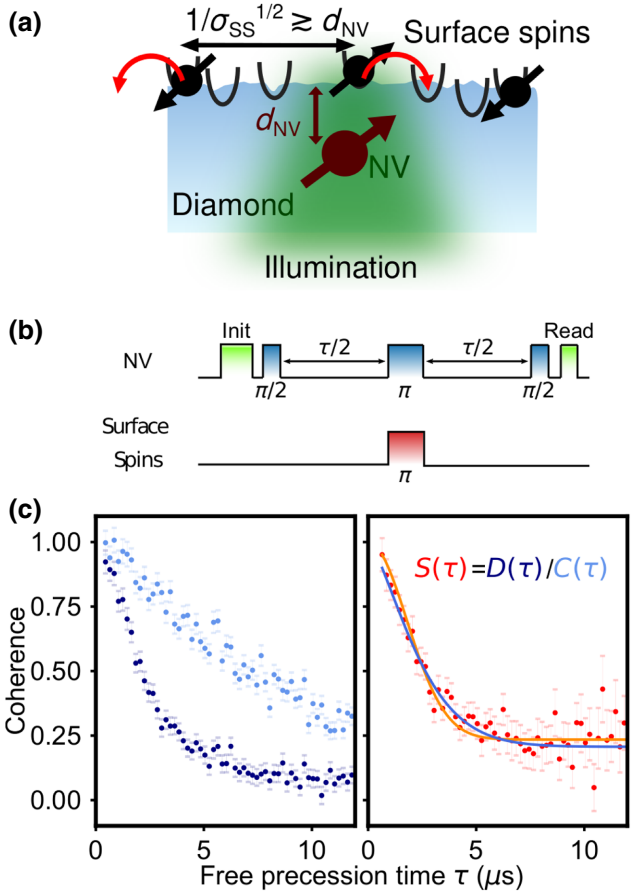


FIG. 1. Double electron-electron resonance of diamond surface spins. (a) Sketch of the experiment, depicting a near surface NV center (dark red arrow) with nearby traps (black wells) that are transiently occupied by spin 1/2 defects (black arrows). A 532-nm laser is used to polarize and readout the NV center. (b) Pulse sequence for DEER, where simultaneous microwave π pulses are applied to the NV center (blue block) and surface spins (red block) midway through the free precession interval τ. Green blocks indicate NV center spin initialization and readout. (c) Spin echo, $C(\tau)$ (light blue), and DEER, $D(\tau)$ (dark blue), curves for a representative NV center with coherence oscillations caused by slight magnetic field misalignment (left panel). In order to investigate the DEER signal $S(\tau)$ (red) without the influence of decoherence from other sources, $D(\tau)$ is divided by the coherence curve $C(\tau)$ (right panel). The orange fit is constrained to a stretching factor $n = 2$, while n is a fit parameter for the blue curve. Error bars in (c) are the standard error of the mean due to photon shot noise, propagated point by point for the right panel.

with no π pulse applied to the surface spins [10]. Since the surface spin contribution to decoherence is not canceled by the echo, the resulting decay is equivalent to the free induction decay (FID) caused by the surface spins alone. This normalization technique has the additional benefit of removing oscillations caused by the periodic entanglement with the ^{13}C nuclear spin bath [15] or slight misalignment between the NV center quantization axis and the

externally applied magnetic field [16]. When such oscillations are not present, we fit the decoherence and DEER signals simultaneously, as described in Appendix A.

We fit the resulting FID to an exponential decay of the form $S(\tau) = e^{-(\Gamma_{\mathrm{DEER}}\tau)^n}$, where τ is the total free evolution time, Γ_{DEER} is the DEER decay rate, and n is the exponential stretching factor. The fitted decay rates for 107 NV centers in nine diamond samples are plotted in Fig. 2(a). For samples with similar processing histories (left plot), the couplings for a given depth cluster together, while samples with different surface preparations (right plot) show different offsets, which we interpret as arising from different densities (σ) of surface spins. Surface preparation details can be found in Appendix A. The lower surface spin densities measured in Fig. 2(a) (right) indicate that there is some natural variation in the surface spin density, and that some surface treatments may be used to lower the surface spin density below what is typical of diamonds receiving a standard surface preparation.

To demonstrate control over the surface spin density, we anneal one diamond sample at 650 °C in vacuum and measure the same NV centers before and after this procedure (Appendix B). In Fig. 2(b), we plot Γ_{DEER} for the same NV centers before and after this annealing procedure and find that, indeed, the DEER decay rates are significantly reduced following annealing, suggesting a reduction in the density of surface spins. We also characterize NV center coherence, without the additional π pulse on the surface spins, before and after annealing and observe that changes in coherence are minimal (Appendix B), suggesting that surface spins are not primarily responsible for the reduction of coherence to near-surface NV centers [4,8,12]. We note that even carefully prepared diamond surfaces host numerous other electronic traps that can be inferred from x-ray spectroscopy methods [10], and that in our samples these other traps are likely dominant sources of electric and magnetic field noise at the surface.

Previous work has treated the surface spin bath as a dense, 2D ensemble of spins with fixed positions but spin states that are initialized randomly between sequences [5,7,12,17]. The DEER decay rate can then be computed as

$$\Gamma_{\mathrm{DEER}} = \left(\frac{\mu_0}{4\pi}\right)\frac{\sqrt{3\pi\sigma}\gamma_e^2\hbar}{8d_{\mathrm{NV}}^2}, \tag{1}$$

where μ_0 is the permeability of free space, γ_e is the electron gyromagnetic ratio in units of radians per second, and $\hbar$ is the reduced Planck constant. This decay arises from an ensemble average over many realizations of spin state configurations, with fixed locations, in which there is no dominant proximal spin. The resulting distribution of total magnetic field amplitudes sensed by the NV center is then normally distributed, which is a requirement for the derivation of Eq. (1) and also leads to $n = 2$ in the exponential decay [18] (Appendix C).

Using this model, we fit DEER decay curves and extract a surface spin density using Eq. (1). However, computing the surface spin density in this way leads to an unphysical inconsistency, namely that the average nearest-neighbor spin-spin separation is greater than the depth of most of the NV centers measured here. The lower limit of the shaded region in Fig. 2(a) is plotted from Eq. (1) for a surface spin density of 0.001 nm^{-2}, or an average nearest-neighbor spin-spin separation of approximately 15 nm (Appendix E). In this regime, the assumption that the NV center sees an average spin density used to derive Eq. (1) is then invalid. For such a low density of spatially fixed surface spins, Eq. (1) should not be valid because the NV center primarily senses the field from the few nearest surface spins and hence the distribution of total magnetic field amplitudes is no longer Gaussian. In this limit, the DEER decay signal should exhibit signatures of coherent coupling between the NV center and one or few surface spins, namely, there should be oscillations in the DEER decay [7]. Additionally, different NV centers should statistically sample different surface spin configurations, leading to a large variation in Γ_{DEER} among NV centers [12]. Instead, we observe that Γ_{DEER} is fairly consistent at similar depths, and in particular does not exhibit increased variance for shallower depths, which is expected for static spins due to the increased likelihood of finding a few dark spins that are markedly closer to the NV center than the others. In the extremely dilute limit, the NV center would primarily sense the field from a single surface spin, which results in strong coherent oscillations rather than an exponential decay. While some of these strongly coupled surface spins have been observed previously [5,7] and are found in some of our nine samples (Appendix D), these events are relatively rare, occurring in $< 2\%$ of measured NV centers. The mismatch between the measured densities of surface spins and the apparent rarity of coherent oscillations has been noted previously [12], but never explained.

III. EXPONENTIAL STRETCHING FACTORS

When Eq. (1) is valid, a Gaussian exponential decay [Fig. 1(c), orange line] is expected, but for many NV centers, a lower value of n produces a more optimal fit [Fig. 1(c), blue line]. In Fig. 2(c), we examine the fitted exponent across all of the samples depicted in Fig. 2(a). For the majority of the shallow NV centers we investigated, the extracted stretching factor is not consistent with Gaussian decay. Furthermore, we observe that shallower NV centers exhibit lower fitted exponents (Pearson's correlation coefficient 0.43, $p < 5 \times 10^{-6}$), and that many NV centers exhibit $n < 1$. The exponential stretching factor n is related to the form of the noise generated by the environment spins [18]. In particular, a Gaussian ($n = 2$) decay occurs for a quasistatic bath, while a smaller exponent may arise from a finite bath correlation time. For a white noise

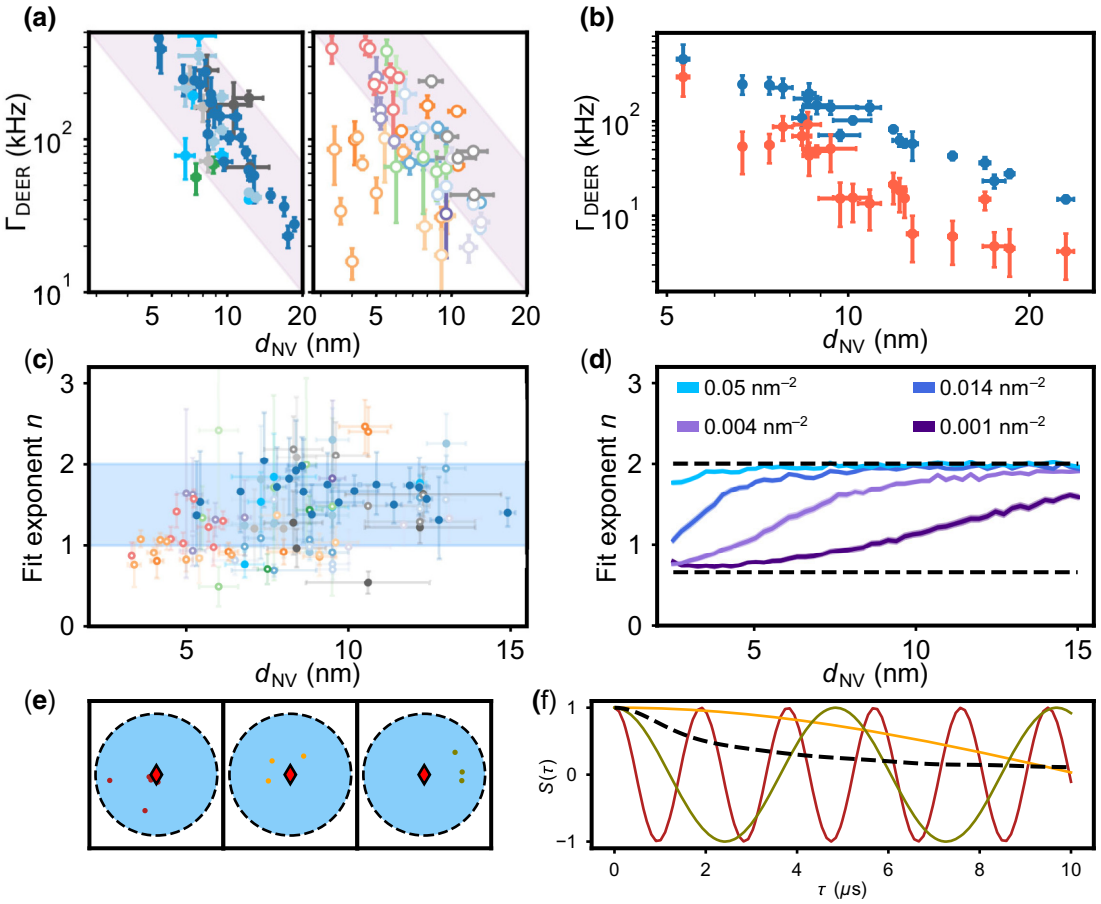


FIG. 2. DEER coupling rates and stretching factors. (a) Fitted DEER couplings as a function of depth for numerous NV centers. Colors represent NV centers measured in a particular sample, while each data point is an individual NV center measurement. All samples in the left plot have nominally identical surface preparations (activated or reset surface, SI), while samples in the right plot represent different surface preparations. Surface preparation details and the sample color key are available in Appendix A. Shaded region indicates expected couplings for surface spin densities of 0.001–0.05 nm^{-2}, according to Eq. (1). The lower limit corresponds to an average nearest-neighbor spin-spin separation of approximately 15 nm, while the upper limit is chosen to encompass the largest Γ_{DEER} shown. (b) DEER coupling versus depth for the same NV centers in Sample 8 before (blue) and after (red) annealing at 650 °C. (c) Fitted exponent as a function of depth for samples in (a). Filled markers are samples with similar surfaces. Shaded region indicates exponents between 1 and 2. (d) Fits of the exponent versus depth for simulated data. Colors indicate different simulation spin densities of 0.05 nm^{-2} (light blue), 0.014 nm^{-2} (dark blue), 0.004 nm^{-2} (light purple), and 0.001 nm^{-2} (dark purple). Shaded regions are confidence intervals for fits of simulated data. Dashed lines indicate limiting values of $n = 2$ and $n = 2/3$. (e) Monte Carlo simulation instances of random spin configurations with spin density 0.005 nm^{-2} (dots) relative to a NV center that is 4 nm deep (red diamond). Plotted spins fall within an 8-nm radius circle of the NV center (dashed circle). (f) DEER signals for the representative configurations of (e), where signals from individual repetitions are dominated by a few nearby spins. Configurational averaging over 10 000 iterations gives rise to exponential decay (dashed line). All error bars are 68% Bayesian credible intervals of the fit parameters.

bath, this exponent saturates at $n = 1$, so the observation of $n < 1$ for many NV centers is surprising. Additionally, measurements of the surface spin T_1 times for a select number of NV centers in this sample (Appendix D) find that surface spin T_1 times are generally > 30 μs, which is long compared to the spin-echo decay time of the shallow NV centers, and so the quasistatic assumption is justified. We also note that, although a short surface spin T_1 time can reduce the value of n from 2 towards 1, the stretching factor as a function of depth should display a trend opposite to that of Fig. 2(c). This is because, for a given surface spin density and relaxation time, a very shallow NV center will experience a full DEER FID collapse before any surface spin relaxation has occurred ($n = 2$), while a deep NV center has a slow enough DEER decay rate that it is able to sample the noise bath created by the fluctuating surface spins ($n < 2$) (Appendix F). In Fig. 2(c), the opposite is observed: the extracted exponent of the DEER decay curve becomes larger for deeper NV centers. Additionally, we note that data points in Fig. 2(c) with low values of the stretching factor are more likely to come from samples with low surface spin densities, as shown in Fig. 2(a).

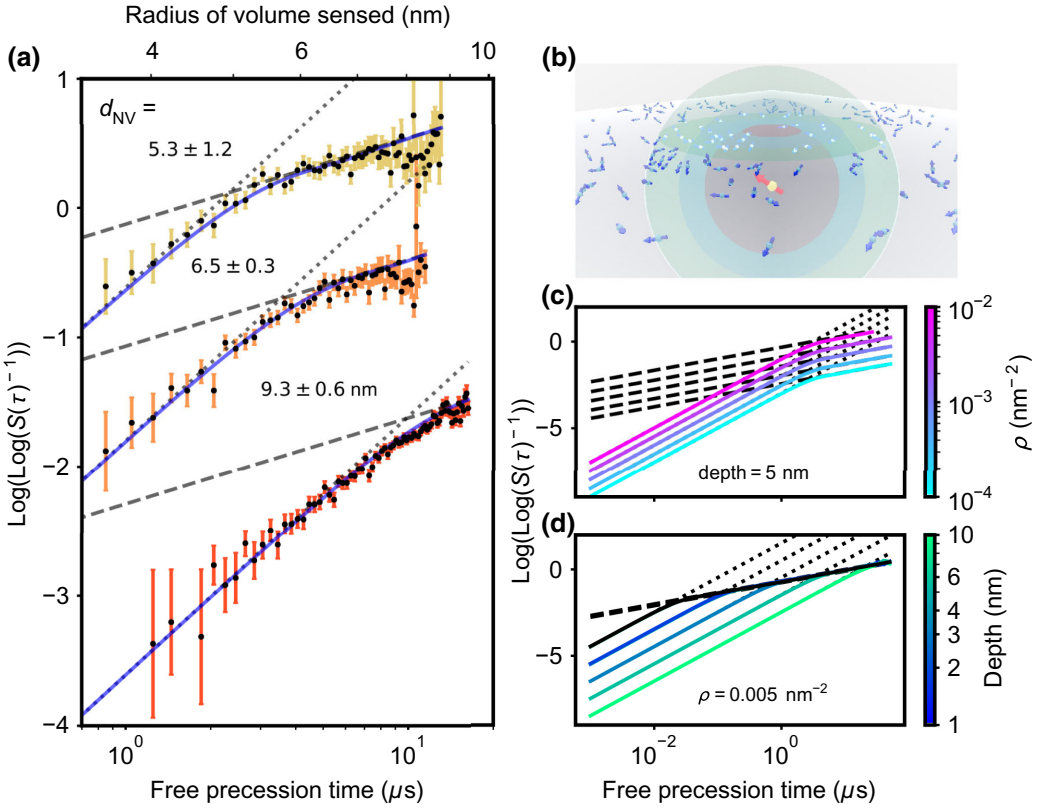


FIG. 3. Temporal transition of the stretching factor. (a) DEER echo signal from shallow NV centers in Sample 9 with three different depths extracted from the proton NMR signal. Error bars are the standard error of the mean due to photon shot noise. The blue line is a fit with the analytical result for configurational averaging (SI). The dashed line is $n = 2$ and the dotted line is $n = 2/3$. Curves are offset for clarity. (b) Diagram depicting the expanding strong dipolar coupling volume of the NV center as the duration of the echo sequence increases. The form of the signal detected by the NV center changes as the volume intersects with the 2D plane of the surface spins. (c) Simulations of the DEER echo signal detected from a 5-nm-deep NV center near a surface with a logarithmically spaced density of 10^{-4} nm^{-2} (blue) to 10^{-2} nm^{-2} (purple) where surface spins undergo configurational averaging. The dashed lines correspond to $n = 2$ and the dotted lines correspond to $n = 2/3$. (d) Simulation of the DEER echo signal detected from NV centers with log-spaced depths from 1 nm (blue) to 10 nm (green) with configurational averaging of the surface spins. The density of simulated surface spins is 0.005 nm^{-2}. The solid black line is the analytic solution for a 1-nm-deep NV center (SI).

Finally, for a few NV centers, we sample the coherence decay with much finer time steps to interrogate the shape of the decay in more detail. In Fig. 3(a) below, we plot the logarithm of the inverse DEER signal, $\log\{S(\tau)^{-1}\}$, as a function of free precession time on a log-log plot so that the stretching factor is given by the slope of the line. Instead of a single slope, we observe a transition in the stretching factor at some critical time, which increases with increasing NV center depth.

IV. CONFIGURATIONAL AVERAGING OF SURFACE SPINS

In order to explain our observations showing the lack of statistical variation of Γ_{DEER} among NV centers at a given depth, the lack of coherent coupling to surface spins, the observed anomalously small decay exponents, and the transition in decay exponents with time, we consider a model wherein most of the surface spins do not have permanently fixed locations, but can occasionally "hop" between unoccupied sites such that the position is stationary for one experimental sequence and may change between sequences, as depicted in Fig. 1(a). After the many experimental repetitions necessary for NV center measurements, the NV center then effectively senses a surface spin ensemble averaged over many positional configurations. For a dilute bath, such an averaging drastically changes the form of the DEER signal due to strong fluctuations of the dipolar coupling strength stemming from the shot-to-shot fluctuations in the spin positions [19,20]. In addition to reducing the variability in DEER decay for very shallow NV centers, stretched exponents arise naturally from sampling different configurations of the spin bath, each with its own characteristic decay time [20,21]. This model differs from normal spin diffusion and is consistent with the long surface spin relaxation times measured

in our samples: the spins can have very long relaxation times during a measurement but suddenly change positions between measurements, realizing a new configuration of quasistatic spins. Other electronic defects in diamond are known to change occupation under the intense illumination of the 532-nm laser used for NV center polarization and readout [22,23], and so it is conceivable that optical initialization and readout are responsible for inducing reconfiguration. For example, occasional repumping of the NV center into its negative charge state at cryogenic temperatures is well known to cause shifts in the electrostatic environment of the defect, and these shifts are thought to come from the population and depletion of nearby charge traps by photoexcited carriers [24,25]. The required averaging (about 10^6 repetitions) in our experiments would make it difficult to determine the frequency of reconfiguration or to observe an effect from illumination power, as even a small probability of hopping results in sampling many spatial distributions of surface spins.

In order to quantitatively investigate our model, we perform Monte Carlo simulations of the DEER signal resulting from configurational averaging of a low density of surface spins, as shown in Fig. 2(d) (insets). By fitting the resulting FID of a large number (10^5) of averages to the same exponential decay function as was used to fit the data in Fig. 2(c), the depth dependence of n is readily apparent and is in qualitative agreement with the trend observed in Fig. 2(c). The fitted values of n approach 2 for large depths [$d_{\mathrm{NV}} > 1/(2\sqrt{\sigma})$], but gradually transition to a value of 2/3 as the depth is decreased below the average surface spin separation.

By integrating over the possible configurations of surface spins [26], we can obtain an analytic expression for the DEER signal measured by a NV center external to the spin bath (Appendix C). For short times such that $\mu_0\gamma_e^2\hbar\tau/4\pi \ll d_{\mathrm{NV}}^3$, we recover Eq. (1) and $n = 2$, even though the spins are assumed to be nonstationary in this case. This result is also the limit of small interspin spacing compared to NV center depth, so the models should be indistinguishable for a dense ensemble of surface spins. In the opposite limit, $\mu_0\gamma_e^2\hbar\tau/4\pi \gg d_{\mathrm{NV}}^3$, we instead find that

$$\langle S(\tau)\rangle_c = \exp\left(-\frac{9\sqrt{\pi}\Gamma(11/6)\sigma}{5}\left(\frac{\mu_0\gamma^2\hbar}{8\pi}\right)^{2/3}\tau^{2/3}\right), \tag{2}$$

where $\Gamma(x)$ is the gamma function. The stretching factor $n = 2/3$ for a 2D bath of spins has been predicted previously [27,28], and its appearance here implies that, for long times, the entire system, including the sensor, behaves as if it is two dimensional.

Unlike in the case of static spins, the DEER signal for configurationally averaged surface spins must transition from an initial Gaussian ($n = 2$) decay into the stretched exponential ($n = 2/3$) decay given by Eq. (2). In our finely sampled data in Fig. 3(a), we find that the data are well fit by our analytical model, which captures the transition regions that move to later times for increasing depths. Because the location of this transition is sensitive to depth alone, it is possible to extract a depth for each NV center from these measurements independent of the surface spin density. These extracted depths are in good agreement with those measured by proton NMR, which additionally indicates that the spins we address are indeed at the diamond surface. NMR measurements, along with additional datasets, are available in Appendix G. In Appendix H, we perform Bayesian model comparison between the static and configurationally averaged models and conclude that the nonstationary model is more consistent with our data.

In Figs. 3(c) and 3(d), we examine the simulated behavior of n in more detail by plotting $\log\{S(\tau)^{-1}\}$. We note a change in the stretching factor from $n = 2$ to $n = 2/3$ at a depth-dependent critical time, similar to what is observed in the data in Fig. 3(a). The shift of this transition to earlier times for shallower NV centers indicates that the intermediate values of n measured in Fig. 2(c) and simulated in Fig. 2(d) are not the result of a continuously varying exponent, but rather of averaging different numbers of points that lie in either the $n = 2$ or $n = 2/3$ regions of the decay and fitting to a single exponent, as is common practice. We emphasize that this does not mean that the value of the stretching factor is depth dependent; only the free precession time at which the transition occurs depends on depth. Fitting the incorrect but commonly used single exponential function to data that in reality contains two different exponents at different times will give rise to the variation with depth observed in Fig. 2(c).

In some cases, repositioning of surface spins can be detected more directly. In Fig. 4(a), we show DEER traces from the same NV center measured minutes apart, showing both a coherent coupling signal that dips below 0 (red), and a weak incoherent coupling signal observed a short time later (orange). By binning into smaller numbers of averages for a fixed free precession interval of 2 μs, it is possible to see the switching behavior in real time in Fig. 4(b). We interpret these results as indicating the presence or absence of a particularly stable proximal spin, and hence as evidence that the occupation of such traps can change. While our model does not rule out the existence of some number of stable, spatially fixed spins, based on the percentage of measured NV centers in our samples that display coherent oscillations, we estimate an upper bound on the number of stationary surface spins of approximately 10% of the total spins detectable by DEER (Appendix I).

The time-dependent behavior of the exponent n can be understood by considering the dipolar coupling between a NV center and the sheet of surface spins at a distance d_{NV} away. Initially, the NV center senses a small contribution from all of the spins, which results in a Gaussian ($n =$

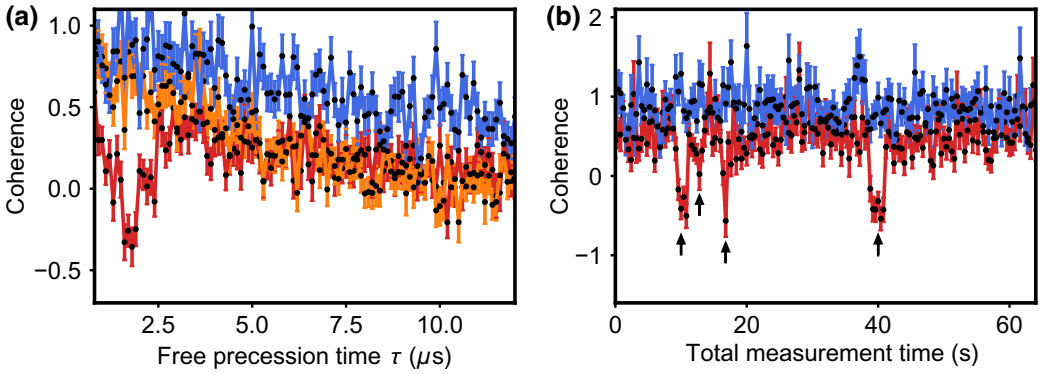


FIG. 4. Transient strong coupling of surface spins. (a) Coherence (blue), coherently coupled DEER signal (red), and incoherent bath signal (orange) from the same NV center measured at different times. (b) Time series for the NV center in (a), with a fixed free precession interval of 2 μs. The DEER signal drops below 0 for short intervals, indicating transient coherent coupling to a nearby dark spin (black arrows). Error bars are the standard error of the mean due to photon shot noise.

2) decay. Once the phase accumulated from the nearest spin reaches a large enough value $[(\mu_0/4\pi)(\gamma_e^2\hbar/d_{\mathrm{NV}}^3)\tau \sim 2\pi]$, further time evolution contributes only an oscillatory phase while more distant spins still contribute a summation of many small phases. This process can be considered with a simple geometric picture. The quantity $[(1/2\pi)(\mu_0/4\pi)\gamma_e^2\hbar\tau]^{1/3} \approx (52\ \mathrm{MHz\,nm^3}\tau)^{1/3}$ defines the radius of a sphere, centered on the NV center, that increases with the free evolution time. Spins within this sphere contribute significant phase ($\gtrsim \pi$) to the NV center. Initially, there are no surface spins within this sphere. The sphere of strong coupling will increase in size until it intersects the surface when $(52\ \mathrm{MHz\,nm^3}\tau)^{1/3} \approx d_{\mathrm{NV}}$, at which time spins begin to interact strongly with the NV center. As is depicted in Fig. 3(b), the number of strongly interacting spins grows with the cross-sectional area that intersects the surface plane, which grows like $\tau^{2/3}$. It is only once the volume of strong interactions grows to include the plane in which the spins dwell that the dimensionality of the bath impacts the NV center coherence, and the NV center decay reflects this dimensionality for long times. In a sense the NV center becomes a member of the 2D bath for sufficiently long interaction times, which is supported by the absence of any dependence on d_{NV} in the long-time limit given by Eq. (2).

We note that all of these observations depend crucially on the use of a single NV center to investigate the spin bath dynamics. Bulk ESR techniques are typically not sensitive enough to detect the extremely low density of spins on the diamond surface, and more importantly the behavior of the FID curve would be qualitatively different. In a bulk sample, the crossover time is set by the minimum possible interspin distance rather than the average interspin distance, so even for a dilute system, the relevant characteristic distance will be of the order of angstroms. This small distance results in a crossover time of about 10 ps so that only the long time $n = 2/3$ behavior would be observed. A separable sensor with a precisely defined position is necessary to delay the crossover time to a measurable duration.

V. CONCLUSION AND OUTLOOK

The inferred dynamics of the surface spins have immediate consequences for surface-spin-based sensing of external targets and for using surface spins as a model system to explore many-body interactions in two dimensions. In particular, the stability of these spins needs to be taken into account and may need to be enhanced. Samples prepared with different processing might host more stable surface spins [7]. Fermi-level engineering using heteroatom doping [29–32], surface transfer doping [33], or different surface chemistry [34] could potentially be employed to create more stable surface spin defects. A more careful survey of coherently coupled surface spins could help elucidate what makes them unique and hence what might be done to stabilize more itinerant defects. For example, the coupling strengths of coherently coupled surface spins observed in this work are consistent with them being at the surface, but there is some evidence that they may reside below the topmost atomic layer (Appendix D). Nano-MRI [5,35] could be used to more carefully localize strongly coupled spins, which could help elucidate factors important in their formation. Understanding under what conditions stable, coherently coupled spins form is the first step towards creating them systematically, perhaps by intentionally introducing lattice defects and passivating the diamond surface. Lacking this, other, potentially more stable systems, such as spinful defects in 2D materials [36,37], could be transferred onto the diamond surface and interrogated by NV centers.

While the observations detailed in this manuscript provide evidence that the surface spin bath undergoes configurational averaging, observing this phenomenon directly is

challenging. A possible candidate for the microscopic origin of surface spins is surface-related traps arising from defects, dangling bonds, and sp^2 carbon, which give rise to energetically shallow states near the conduction band minimum [10]. Such defects would be ionized by the relevant excitation wavelengths used to address NV centers. The low readout fidelity of room-temperature NV center experiments make correlations between sequences difficult to discern, and so the effects of a small reconfiguration probability are indistinguishable from a bath that reconfigures at each repetition. A more accurate determination of the reconfiguration probability and mechanism could be made with single-shot readout of the NV center through resonant excitation [38,39] or mapping the electronic spin state to the nitrogen nuclear spin [6]. If photoexcitation is indeed the mechanism that drives surface spin reconfiguration then nonoptical techniques such as magnetic resonance force microscopy [40] would be valuable tools since they avoid optical readout entirely. Using the framework outlined here, spin ensembles on diamond surfaces and in other materials can be studied with scanning NV center experiments [41,42] where the distance to the sample, and hence the exponential stretching factor transition time, can be tuned continuously.

More generally, the techniques for understanding surface environment dynamics presented here may be broadly applicable in other quantum systems. The dynamics of proximal probes can carry significant information: the minimum impurity separation, the dimensionality of the environment, and the scales of finite dimensions are all reflected in the behavior of the decay, constituting a promising tool for studying the disordered environment of quantum systems.

The unprocessed datasets and simulation code analyzed during the current study are available from the corresponding author on reasonable request.

ACKNOWLEDGMENTS

We thank Ania Jayich, Norman Yao, Soonwon Choi, Alastair Stacey, Cherno Jaye, and Andrew Evans for fruitful discussions. Undergraduate researchers Trisha Madhavan and Rohith Karur contributed to surface spin measurements that helped us to calibrate the annealing procedure. This work is supported in part by the DARPA DRINQS program (Grants No. D18AC00015 and No. D18AC00033), the Center for Ultracold Atoms (NSF PHY-1734011), and the Moore Foundation (Grant No. 4342.01). Sample surface preparation and spectroscopy was partially supported by the NSF under the CAREER program (Grant No. DMR-1752047) and through the PCCM (Grant No. DMR-1420541). L.V.H.R. acknowledges support from the DOD through the NDSEG Fellowship Program. E.K.U. acknowledges support from the NSFGRFP (Grant No. DGE1144152). D.B. acknowledges support from the NSFGRFP (Grant No. DGE1745303) and the Hertz Foundation. S.S. acknowledges support from the PMU-B (Grant No. B05F630108). J.C. acknowledges financial support from the MST, Taiwan (MOST-109-2112-M-018-008-MY3). M.F. is supported by an appointment to the IC Postdoc Research Program by ORISE through an interagency agreement between the U.S. DOE and the ODNI. A.G. acknowledges support from the NKFIH in Hungary for the Quantum Technology Program (Grant No. 2017-1.2.1-NKP-2017-00001), the National Excellence Program (Grant No. KKP129866), the Quantum Information National Laboratory, and the EU QuantERA II MAESTRO project and from the European Commission for the ASTERIQS and QuMicro projects (Grant No. 820394 and No. 101046911). V.V.D. is partially supported by QuTech Physics Funding (QTECH, program 172, No. 16QTECH02), which is partly financed by the NWO. This work was performed in part at the Imaging and Analysis Center at Princeton, the Center for Nanoscale Systems at Harvard (NSF Grant No. 1541959), and the research beam line U7A of the National Synchrotron Light Source operated for the DOE by Brookhaven National Laboratory (Contract DE-AC02-98CH10886).

Note added.— Recently, we became aware of complementary work [43] using probe spins to measure the many-body noise from a strongly interacting dipolar system. Both works use the decoherence profile of the probe spin to characterize the dimensionality and dynamics of the many-body system.

APPENDIX A: METHODS

1. Sample processing

All the samples measured for this work are electronic grade samples (< 5 ppb nitrogen, < 1 ppb boron) from Element 6. Those with an additional overgrown ^{12}C enriched layer are indicated in Table I. Most of the samples have surfaces that are prepared prior to implantation following the procedure outlined in Ref. [10]. Briefly, the preimplantation surface preparation begins with scaife polishing, followed by etching in Ar/Cl_2 followed by O_2 plasma, and then annealing to 1200 °C in a vacuum tube furnace to graphitize and remove the first few nanometers of the surface. Sample 3 and Sample 9 do not undergo this surface preparation prior to implantation and are implanted with the "as grown" surface. All samples are implanted with ^{15}N, and annealed to 800 °C in a tube-vacuum furnace to form NV centers.

After the activation anneal, NV centers are measured under a variety of different surfaces, as noted in Table I. NV centers are randomly selected and included in the dataset provided they exhibit enough ESR contrast to be measured. NV centers with very short coherence times ($\Gamma_{DEER} < 1/2T_2$) are not included, as insufficient information is present to confidently extract fit parameters

TABLE I. Diamond samples.

Sample	C12 enriched	Preimplant surface prep	Color	Surface condition
Sample 1		×		Activated
				Oxygen annealed
				Reset
				Oxygen annealed
Sample 2		×		Oxygen annealed
Sample 3	×			Activated
				500 °C annealed
				650 °C annealed
Sample 4		×		Reset
Sample 5	×	×		Oxygen annealed
				Reset
Sample 6		×		Oxygen annealed
				Reset
				500 °C annealed
Sample 7		×		Activated
Sample 8	×	×		Activated
Sample 9	×			Oxygen annealed

of the DEER signal. In some cases, DEER coupling is characterized in the same sample after different surface processing steps. The final surface conditions noted in Table I are prepared as follows.

(a) Activated: NV centers are measured directly after 800 °C activation anneal and triacid cleaning (reflux sample in 1:1:1 sulfuric, nitric, and perchloric acids for 2 h).
(b) Oxygen annealed: NV centers are measured after heating the sample to 440–460 °C while flowing an atmosphere of oxygen gas over the samples, followed by cleaning in a 2:1 piranha mixture (sulfuric acid and hydrogen peroxide) for 20 min.
(c) Reset: NV centers are measured after 'resetting' the surface to a state comparable to the activated surface by reannealing the sample to 800 °C in a vacuum tube furnace at around 10^{-6} mbar, and subsequently triacid cleaning the sample.
(d) 500 °C annealed: NV centers are measured after annealing the sample to 500 °C in a vacuum tube furnace, followed by triacid cleaning.

2. Measurement setup

The details of our measurement setup have been published elsewhere [10], but we briefly review them here for completeness. The data for this manuscript are taken on several measurement setups that are constructed in a similar manner but with different components.

NV center measurements are performed on a home-built confocal microscope. NV centers are excited by a 532-nm frequency doubled Nd:YAG solid-state laser that is modulated with an acousto-optic modulator. The beam is scanned using galvo mirrors and projected into an oil immersion objective (Nikon, Plan Fluor 100X, NA = 1.30) with a telescope in a 4f configuration. Laser power at the back of the objective is kept between 60–100 μW, approximately 25% of the saturation power of a single NV center, in order to avoid irreversible photobleaching. A dichroic beamsplitter separates the excitation and collection pathways, and fluorescence is measured using a fiber-coupled avalanche photodiode (Excelitas SPCM-AQRH-44-FC). A neodymium permanent magnet is used to introduce a dc magnetic field for Zeeman splitting, and the orientation of the magnetic field is aligned to within 1° of the NV center axis using a goniometer. Measurements are performed at moderate magnetic fields of approximately 300 Gauss so that the surface spin and NV center transitions are frequency resolved.

Spin manipulation on the NV center and surface spins is accomplished using microwaves. Two signal generators, which output the NV center and DEER transition frequencies, are separately gated with fast single-pole double-throw switches (Mini-Circuits ZASWA-2-50DR+). In one setup, these signals are then amplified by Mini-Circuits ZHL-16W-43+ and ZHL-100W-13+, respectively, before being combined with a high-power resistive combiner and delivered to the sample via a coplanar stripline. In another setup, the gated signals are combined with a resistive combiner (Mini-Circuits ZFRSC-42-S+) and amplified with a high-power amplifier (Ophir 5022A) before being delivered to the sample via a coplanar stripline. The stripline is fabricated by depositing 10 nm Ti, 1000 nm Cu, and 200 nm Au on a microscope coverslip. Following metallization, the stripline is photolithographically defined and etched with gold etchant and hydrofluoric acid. Finally, a 100-nm layer of Al_2O_3 is deposited on top of the fabricated stripline via atomic layer deposition (ALD) to protect the metal layer. Pulse timing is controlled with a Spincore

PulseBlaster ESR-PRO500 with 2-ns timing resolution. The DEER pulse is either applied simultaneously to the NV center pulse, as depicted in Fig. 1 of the main text, or it is offset from the NV center pulse in order to avoid saturation of the amplifier in some cases. In the latter case, an identical DEER pulse is applied immediately after the first $\pi/2$ pulse on the NV center as in Ref. [11].

3. Fitting procedures

For the majority of NVs studied, the coherence and DEER signals are fitted simultaneously using Markov chain Monte Carlo (MCMC) using two likelihood functions, where the DEER signal is assumed to arise from a pure DEER decay multiplied by a decoherence envelope. The fit functions for both the DEER and coherence signals are assumed to be of the form $y = a\exp(-(b\tau)^n) + d$, where uniform priors are used for each fit parameter. The fit parameters are estimated to be the means of the resulting posterior distributions of the fit parameters, and the reported errors are the 68% Bayesian credible intervals, which are similar to one standard error. For NV centers that are coupled to ^{13}C atoms and exhibit coherence oscillations, the DEER signal is instead divided by the coherence point by point and the errors of the signal are propagated using Gaussian error propagation before fitting with a single likelihood function. The data in Fig. 3(a) are obtained in the same manner, but the free precession intervals are highly sampled and the data averaged for a long enough period of time that the change in stretching factor is observable. After point-by-point division by the coherence, we take $\log\{S(\tau_i)^{-1}\}$ for each point, and plot on a log-log scale such that the exponential stretching factor is transformed into a linear slope. The resulting datasets are fit using the general configurational averaging model that is derived in Appendix C.

APPENDIX B: REMOVAL OF SURFACE SPINS THROUGH ANNEALING

1. X-ray spectroscopy

To investigate the microscopic origin of surface spins and to develop a procedure for their selective removal, we interrogate diamond surfaces using x-ray spectroscopy. In near-edge x-ray absorption fine structure (NEXAFS) spectroscopy, monochromatic incident x-rays excite core

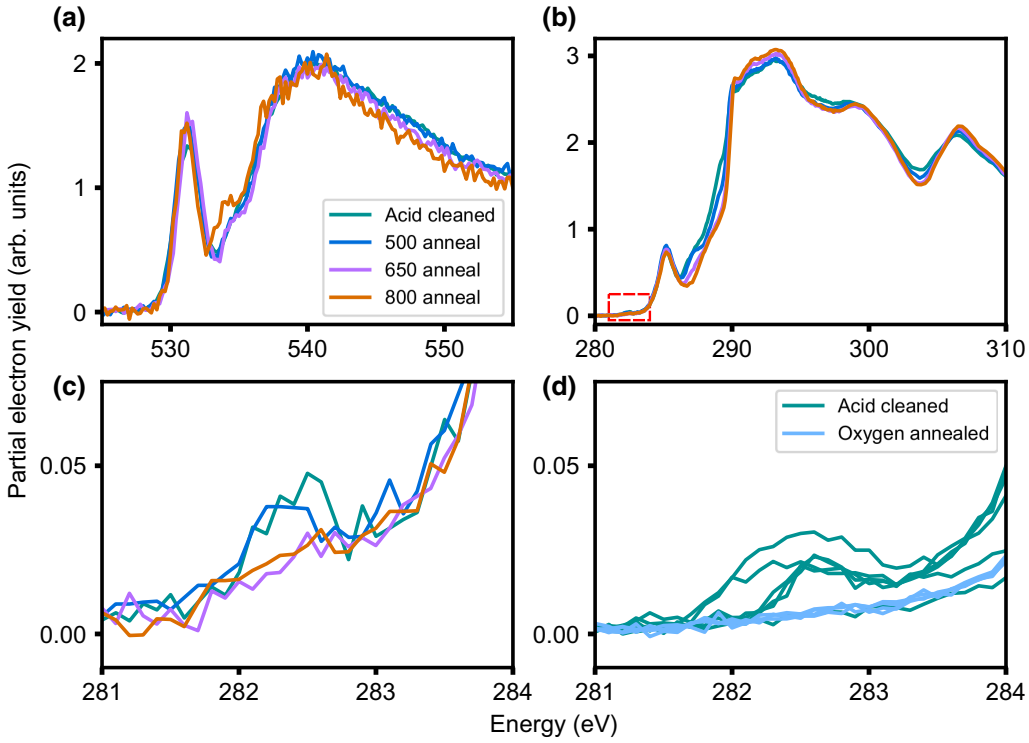


FIG. 5. NEXAFS through annealing. (a) Oxygen edge NEXAFS for a diamond sample with initial acid cleaned surface (green), after 500 °C annealing (blue), 650 °C annealing (purple), and 800 °C annealing (orange). (b) Carbon edge NEXAFS of the same annealing sequence on the same sample. Red box indicates the location of the dangling bond peak, shown in detail in (c). (c) Detailed view of the dangling bond peak showing removal at higher annealing temperatures. (d) Detailed view of the dangling bond peak for several samples with initially prepared acid cleaned surfaces (green) and oxygen annealed surfaces (blue), showing an absence of the peak for oxygen annealed surfaces. Oxygen annealing procedure and suppression of the DEER signal are described in Ref. [10].

electrons, and the electron yield is measured as a function of x-ray energy, giving a signal that is proportional to the unoccupied density of states near the surface [44]. We perform NEXAFS spectroscopy on diamond samples prepared with the standard surface preparation procedure described in Appendix A 1, without implanting NV centers. NEXAFS spectroscopy at the carbon edge reveals a peak at 282.5 eV [Fig. 5(b)]. Previous studies assign this peak to sp^3 dangling bonds at the diamond surface [45]. The presence of this dangling bond peak in the NEXAFS spectra correlates with the surface spin DEER signal measured via shallow NV centers. Specifically, we observe that the peak at 282.5 eV is present in thermally annealed and acid cleaned samples, while it is absent in samples after oxygen annealing, as described in Ref. [10]. Quantitatively, oxygen annealing diminishes this peak from a signal-to-noise ratio of around four to a value below the noise floor [Fig. 5(d)]. Correspondingly, we have also previously shown that individual NV centers show decreased DEER coupling rates after oxygen annealing, and increased coupling after a thermal surface reset process [10].

The correlation between the 282.5 eV peak in NEXAFS and DEER coupling strengths across samples suggests that dangling bonds at the diamond surface may be responsible for the surface spin signal detected in NV center-based sensing experiments. However, there are a variety of observable changes at the diamond surface upon oxygen annealing [10], and a definitive link requires a process for selectively removing dangling bonds. When the samples are annealed in high vacuum at 650 °C for 1 h, the dangling bond peak falls below the noise floor [Fig. 5(c)], but the oxygen edge NEXAFS spectrum is essentially unchanged [Fig. 5(a)]. Thus, annealing at 650 °C leads to selective removal of the dangling bond peak. This annealing procedure also leads to a decrease in DEER coupling for the same NV centers, as shown in the main text.

2. Annealing procedure

Sample 8 is annealed in a vacuum tube furnace with a base pressure around 10^{-6} mbar. The sample is heated from room temperature to 100 °C over 1 h, held at 100 °C for 11 h, heated to 650 °C over 20 h, held at 650 °C for 2 h, then cooled to room temperature.

We note that we have tried annealing in different vacuum systems and have sometimes measured a reduction in DEER coupling at inconsistent temperatures. We attribute these discrepancies to differences in annealing duration, temperature calibration, or thermal anchoring of sample to heater. Other samples show a persistent DEER signal that is not reduced by vacuum thermal annealing or oxygen annealing. Many of these samples either do not go through processing to reduce subsurface damage prior to implantation or have rough surface morphologies from iterative processing.

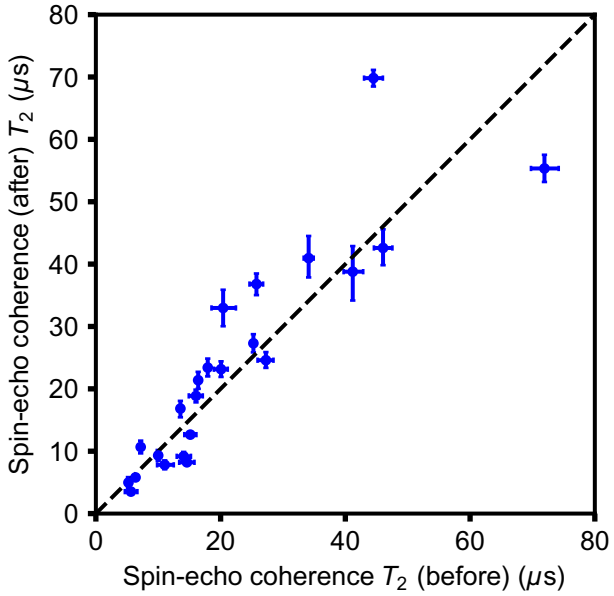


FIG. 6. Coherence times of NV centers in Sample 8 after annealing, plotted as a function of preannealing coherence times. Dashed line has a slope of one.

3. Coherence change with annealing

For the NV centers depicted in Fig. 2(a) in the main text, we also measure the spin-echo coherence time (T_2) of each NV center before and after annealing. The results are plotted in Fig. 6. In contrast to oxygen annealing coherence times of these NV centers are not significantly affected by the vacuum annealing procedure [10]. Because the DEER couplings for these NV centers are reduced by annealing, we interpret these results together to imply that the surface spin bath is not primarily responsible for limiting the coherence time of near surface NV centers in this sample.

4. Spectral decomposition through annealing

The spectral density of the noise bath $S(\omega)$ can be probed by dynamical decoupling [8–10]. We use the non-symmetric XY8 pulse sequence [46] with multiple repetitions to probe different regions of the noise spectrum.

In general, the coherence decay, $C(T)$, of a NV center is given by

$$C(T) = \exp[-\chi(T)], \tag{B1}$$

where T is the total free precession time ($T = N\tau$ for N precession intervals of length τ) and $\chi(T)$ is given by

$$\chi(T) = \frac{1}{\pi}\int_0^\infty S(\omega)\frac{F_N(\omega T)}{\omega^2}d\omega, \tag{B2}$$

where $S(\omega)$ is the magnetic noise spectrum and $F_N(\omega T)$ is the N-pulse filter function given by

$$F_N(\omega T) = 8\sin^4\left(\frac{\omega T}{4N}\right)\frac{\sin^2(\omega T/2)}{\cos^2(\omega T/2N)}. \quad \text{(B3)}$$

When $F_N(\omega T)$ is sharply peaked, Eq. (B2) can be approximated as

$$\chi(T) \approx \frac{TS(\omega)}{\pi}. \quad \text{(B4)}$$

We apply this procedure to the decoherence curves of select NV centers in Sample 8 before and after a 650 °C vacuum anneal to investigate the change in the noise spectrum due to annealing, and find that it is not significantly affected by the annealing procedure. These results are presented in Fig. 7. NV centers in this sample display reduced DEER signals after annealing, and so the absence of change in the noise spectrum indicates that the surface spins do not contribute significantly to the surface noise bath in this sample at the frequencies probed.

5. Atomistic model for surface spins on a (100) diamond surface

X-ray spectroscopy data indicate that carbon sp^3 dangling bonds are a likely source of surface spins (main text, Appendix B). This is a surprising result as surface sp^3 dangling bonds appear to be chemically stable in ambient conditions, but are expected to be highly reactive in atmosphere. In order to set up a plausible model, we assume that some structural disorder is left on the surface after oxygenation, so the surface is not atomically smooth.

We propose a model to explain the existence and chemical stability of surface spins on the (100) diamond surface as sp^3 dangling bonds at a (111) step edge. A (111) surface is essential to create a single dangling bond as, in contrast to a (100) surface, each carbon atom under a (111) surface possesses three bonds downward and one bond upward, from which it is possible to form a single sp^3 dangling bond defect with minimal change of the diamond lattice. In this model, all carbon atom dangling bonds are saturated with OH and the total spin of these systems is $S = 1/2$. We also propose that the (111) facet, which is naturally generated at a step edge of (100) crystalline diamond surfaces, can sterically protect these spinful defects.

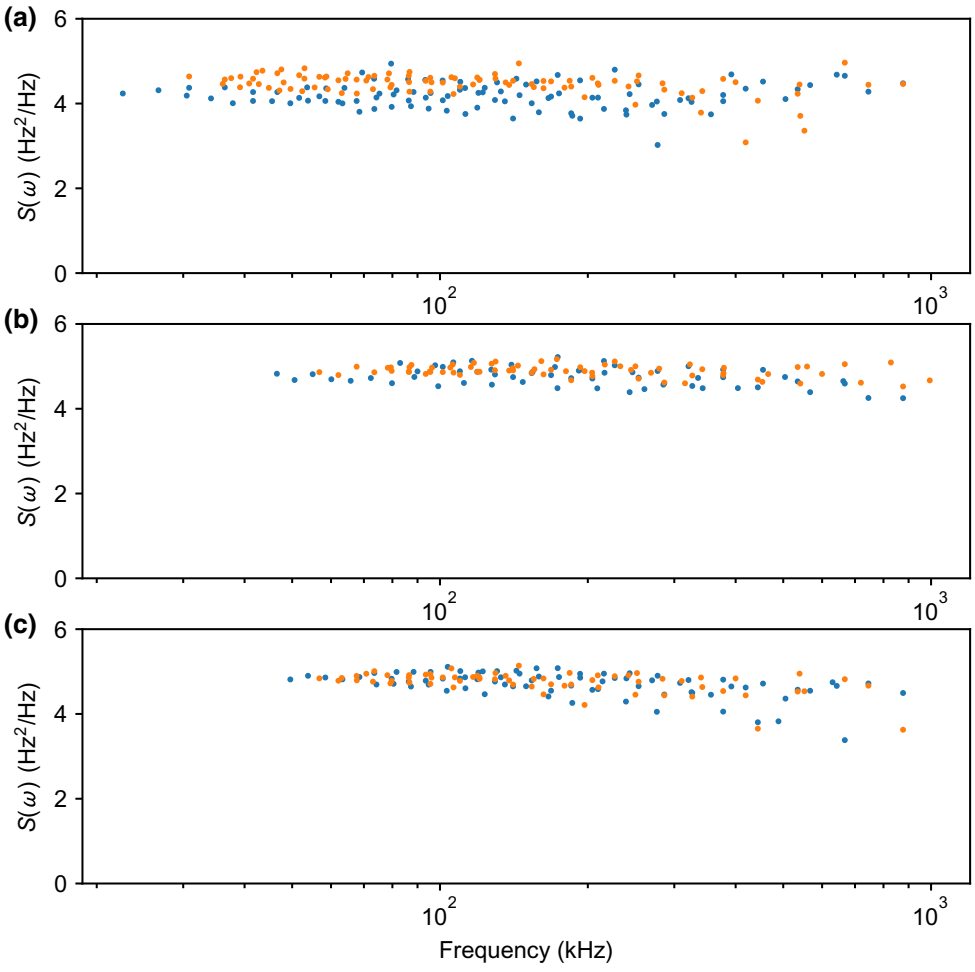


FIG. 7. Spectral decomposition for a sample of NV centers in Sample 8 before (blue) and after (orange) 650 °C vacuum anneal. Depths of NV centers are (a) 10.25 nm, (b) 9.72 nm, (c) 10.42 nm.

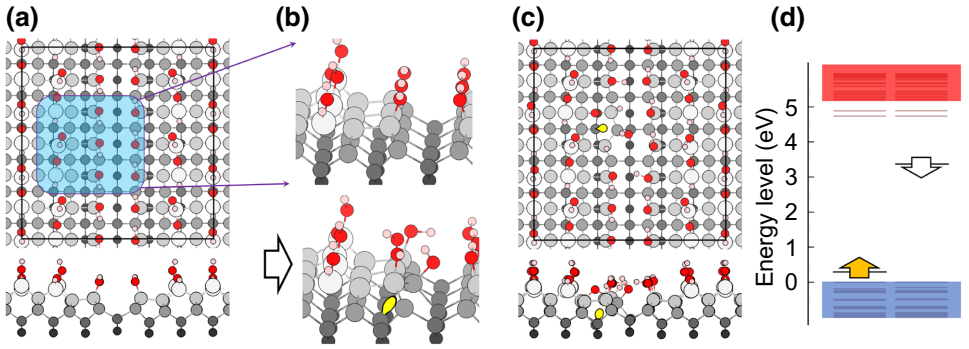


FIG. 8. Surface spin model. (a) The trench structure of a (100)-6 × 6 diamond surface. Surface termination is O/H/OH type, as in Ref. [34], and step carbon atoms are saturated by OH. The red and pink balls are hydrogen and oxygen atoms, respectively. The gray balls are carbon atoms, with the darker color and smaller size indicating deeper positions away from the surface. For the sake of clarity, only the first five layers are presented in the structure plots. (b) Illustration of the formation of a surface spin at the step edge. The yellow lobe represents the surface spin. (c) A plausible structure model of a surface spin on the trench of the diamond (100) surface. (d) The energy-level plot of the modeled surface spin. The valence and conduction bands are depicted as blue and red regions, respectively. The orange and white arrows represent spin-up and spin-down channels, respectively. The spin-polarization energy is high for the dangling bond. We note that the single-particle Kohn-Sham levels do not directly refer to the ionization energies of the dangling bond. The surface states are located at 0.3 eV below the conduction band minimum. The valence band maximum positions are aligned to zero.

It has previously been reported that O/H/OH mixed termination of (100) diamond minimally introduces defect levels to the band gap [34], so we use this surface as our starting model. In order to mimic the local disorder at the surface, we create two step edges forming a trench configuration [47], as shown in Fig. 8(a). A previous study [48] indicated that such a trench structure can be energetically more favorable than the flat structure on the diamond (100) surface. We carry out first-principles calculations on this surface model using the plane-wave based Vienna *ab initio* simulation package [49]. The projector augmented-wave [50] method is used to represent the electron-ionic core interactions. For the exchange-correlation functionals, we use the spin-polarized gradient-corrected Perdew-Burke-Ernzerhof functional [51] for structure optimization and the Heyd-Scuseria-Ernzerhof [52] hybrid functional for electronic structure calculations that would be capable of providing correct defect levels and defect-related electronic transition within 0.1 eV to experiments [53]. A cutoff energy of 370 eV resulted in an equilibrium lattice parameter of diamond of 3.570 Å, which agrees with the experimental value of 3.567 Å. We use a 6 × 6 supercell to simulate the trench models. The thickness of the vacuum layer is more than 10 Å. There are eleven carbon layers; the surface seven layers are allowed to be fully relaxed until the forces are below 0.01 eV/Å and the bottom four layers are fixed at their bulk positions. The k-point sampling is Γ only, which is sufficient to map the Brillouin zone.

In the starting model, a single carbon sp^3 dangling bond is introduced as the hypothetical source of the $S = 1/2$ spin defect. The carbon atoms at the atomic step are more reactive because they do not possess an ideal bonding configuration, and thus may be attacked by a radical such as −OH groups present in the oxygenation process. We simulate this situation as shown in Fig. 8(b). First-principles calculations show that the surface carbon atom attacked by the −OH group spontaneously breaks a bond with a neighbor carbon atom beneath the surface, creating a sp^3 dangling bond with $S = 1/2$ spin as respectively shown in Figs. 8(c) and 8(d). This near-surface carbon sp^3 dangling bond resides in the fourth atomic layer referred to the topmost carbon atom layer that is topographically protected by the direct interaction with species from the gas phase. As expected, the sp^3 dangling bonds are electrically and optically active; the latter may be achieved by photoionization to the negative charge state. The actual photoionization energy threshold may depend on the interaction with other defects at the surface or beneath the surface.

APPENDIX C: DERIVATION OF THE DEER SIGNAL FOR STATIC AND HOPPING SURFACE SPINS

1. System description

We consider a single NV center coupled to a 2D spin bath. We assume the 2D spin bath to be located at $z = 0$, in the $\hat{x}$-$\hat{y}$ plane, and the NV center to be located at $z = -d_{\mathrm{NV}}$. Consistent with a global magnetic field applied along the NV center quantization axis and for [100] diamond, we assume that all spins are oriented in a direction 54.7° from the $\hat{z}$ axis. The Hamiltonian of the system with NV center spin S with resonant frequency ω_s, and surface spins I_i with

Larmor frequencies ω_i can be written as

$$\mathbf{H} = \omega_S \mathbf{S}^z + \sum_i \omega_i \mathbf{I}_i^z + \sum_i J_i \mathbf{S}^z \mathbf{I}_i^z, \tag{C1}$$

where we have assumed that the interaction between the NV center and surface spins takes an Ising form, since they have different resonance frequencies. The interaction strengths J_i are assumed to be dipolar interactions, and we explicitly calculate their distributions below.

In the above equation, we have also neglected the interactions between the surface spins, which may be justified for two reasons: first, local disorder fields experienced by the surface spins that cause their resonance frequencies to vary (see Appendix D, as well as Ref. [5]), and this will suppress spin-exchange processes; second, the Ising interaction component will not change the I^z component of each spin, and thus it will not affect the NV center spin signal. Note that we have also assumed a quasistatic bath for now and neglect the finite T_1 relaxation time of the spins. Spin-exchange processes between bath spins could be approximately treated in a similar fashion as a finite T_1 time. For completeness, we present some results from simulations that include interactions in Appendix F.

We now calculate the DEER signal expected for a NV center in the case of static surface spins and of surface spins undergoing configurational averaging.

2. Derivation of DEER coupling for dense static spins

The combined NV center-surface spin Hamiltonian is given by Eq. (C1). In the rotating frame [54], this simplifies to

$$\mathbf{H} = \gamma_e \mathbf{S}^z B_z^{\mathrm{DEER}}, \tag{C2}$$

where γ_e is the electron gyromagnetic ratio and B_z^{DEER} is the total z component of the magnetic field due to the surface spins at the location of the NV center. We now calculate the signal measured by a NV center during a DEER spin-echo pulse sequence, in the case that other decoherence mechanisms affecting the NV center are negligible.

After spin initialization, the NV center spin starts out in the spin state $|0\rangle$ ($m_s = 0$), and the surface spins have an initial state vector $|I\rangle = |I_1, I_2, \ldots, I_n\rangle$, where $\mathbf{I}_i^z |I_i\rangle = \pm\hbar/2 \, |I_i\rangle$. The initial spin states are random for each repetition of the experiment. Then we apply a $\pi/2$ pulse to the NV center to put the system in the superposition state

$$|\psi_{\mathrm{NV}}(t=0)\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |I\rangle \, . \tag{C3}$$

The NV center spin then evolves for a time $\tau/2$ under the influence of the magnetic field produced by the surface spins. This field is given by the summation of all dipole fields from the n spins:

$$B_z^{\mathrm{DEER}}(t) = \frac{\mu_0}{4\pi} \sum_i \frac{3u_z^i(\vec{m}^i \cdot \vec{u}^i) - m_z^i}{|r^i|^3}. \tag{C4}$$

Here $\vec{u}$ is a unit vector from the NV center to spin i, r^i is the distance to the ith spin, and $m^i = -\gamma_e \hbar \vec{\sigma}^i/2$, and σ^i is the Pauli matrix for spin i.

After a time $\tau/2$, the NV center spin will be in a state

$$\left|\psi_{\mathrm{NV}}\left(t=\frac{\tau}{2}\right)\right\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |I\rangle) + \frac{1}{\sqrt{2}}(|1\rangle \otimes e^{-i\sum_i \phi_1^i} |I\rangle), \tag{C5}$$

where

$$\phi_1^i = -\gamma_e^2 \hbar \frac{\mu_0}{8\pi\hbar} \frac{1}{|r^i|^3}(3u_z^i u_z^i - 1) I_i \frac{\tau}{2} \tag{C6}$$

and I_i is either $+1$ or -1 for each spin, but it is initialized randomly for each experiment.

Next we apply a resonant π pulse to the NV center and surface spins simultaneously, which has the effect of swapping $|0\rangle \leftrightarrow |1\rangle$ and also swapping $|I_i\rangle \rightarrow |-I_i\rangle$. The state is now

$$\left|\psi_{\mathrm{NV}}\left(t=\frac{\tau}{2}\right)\right\rangle = \frac{1}{\sqrt{2}}(|1\rangle \otimes |-I\rangle) + \frac{1}{\sqrt{2}}(|0\rangle \otimes e^{-i\sum_i \phi_1^i} |-I\rangle). \tag{C7}$$

After another free evolution period of $\tau/2$, we have

$$|\psi_{\mathrm{NV}}(t=\tau)\rangle = \frac{1}{\sqrt{2}}(|1\rangle \otimes e^{-i\sum_i \phi_2^i} |-I\rangle) + \frac{1}{\sqrt{2}}(|0\rangle \otimes e^{-i\sum_i \phi_1^i} |-I\rangle), \tag{C8}$$

where

$$\phi_2^i = +\gamma_e^2 \hbar \frac{\mu_0}{8\pi\hbar} \frac{1}{|r^i|^3}(3u_z^i u_z^i - 1) I_i \frac{\tau}{2} = -\phi_1^i. \tag{C9}$$

The final $\pi/2$ pulse on the NV center leads to

$$|\psi_{\mathrm{NV}}(t=\tau)\rangle = \frac{1}{2}(|1\rangle - |0\rangle) \otimes e^{-i\sum_i \phi_2^i} |-I\rangle + \frac{1}{2}(|0\rangle + |1\rangle) \otimes e^{-i\sum_i \phi_1^i} |-I\rangle \, . \tag{C10}$$

Measuring in the $|0\rangle$ basis then gives

$$|\langle 0 \mid \psi_{\mathrm{NV}}\rangle|^2 = |-\frac{1}{2}e^{-i\sum_i \phi_2^i} + 1/2e^{-i\sum_i \phi_1^i}|^2 = \frac{1}{2}\left(1+\cos\sum_i \Delta\phi^i\right), \tag{C11}$$

where

$$\Delta\phi^i = \gamma_e^2\hbar\frac{\mu_0}{8\pi\hbar}\frac{1}{|r^i|^3}(3u_z^i u_z^i - 1)I_i\tau. \tag{C12}$$

If there is only one nearby dark spin coupled to the NV center, we see full contrast oscillations at the coupling frequency of the surface spin.

The more typical situation is one in which the NV center senses many dark spins. At each experimental repetition, the NV center senses a different total magnetic field depending on the random spin states of the surface spins. The total signal is the average of this Bernoulli random process for many trials. For a large number of trials, this can be approximated by a Gaussian process, e.g., $\langle\cos\phi\rangle = e^{-\langle\phi^2\rangle/2}$. This will only be true if the NV center is coupled to sufficiently many spins with comparable coupling strengths, and in the limit of a dilute spin bath the Gaussian approximation will fail because of strong trial-to-trial fluctuations in the strength of dipolar coupling between the NV center and the dark spins, which lead to a distinctly non-Gaussian distribution of acquired phases [19,20].

Computing the average of squares we find that

$$\langle\Delta\phi^2\rangle = \left(\frac{\tau\gamma_e^2\mu_0\hbar}{8\pi}\right)^2\sum_i\frac{1}{|r^i|^6}(3u_z^i u_z^i - 1)^2. \tag{C13}$$

If the surface spins are dense enough that the average spin-spin spacing is much smaller than the NV center depth, d_{NV}, the sum can be approximated as an integral over a continuous density of surface spins, σ. We can write $r = \sqrt{a^2+d_{\mathrm{NV}}^2}$ and $u_z = (1/r)(d_{\mathrm{NV}}\cos\theta + a\sin\theta)\cos\alpha$ and calculate

$$\sum_i\frac{1}{|r^i|^6}(3u_z^i u_z^i - 1)^2 \approx \int_0^{2\pi}\int_0^{\infty}\frac{dad\alpha a}{(d_{\mathrm{NV}}^2+a^2)^3}\left(\frac{3}{a^2+d_{\mathrm{NV}}^2}(d_{\mathrm{NV}}\cos\theta + a\sin\theta)^2\cos^2\alpha - 1\right)^2 = \frac{3\pi}{8d_{\mathrm{NV}}^4} \tag{C14}$$

for $\theta = 54.7^\circ$, the angle of NV centers in (100) diamond with respect to the surface. Converting the population to the coherence, we finally obtain

$$\langle S(\tau)\rangle = \exp\left(-\left(\frac{\mu_0}{4\pi}\right)^2\frac{3\pi\gamma_e^4\hbar^2\sigma}{64d_{\mathrm{NV}}^4}\tau^2\right). \tag{C15}$$

3. Calculation of NV coherence decay for configurationally averaged surface spins

Based on experimental observations of a small stretched exponent and the absence of strongly coupled spins, we calculate the case where surface spins can change positions between measurements, effectively allowing a configurational averaging when evaluating the coherence decay.

Since configurational averaging can result from random spin hopping, we can make use of the methods in Refs. [27,55] to calculate the coherence decay. The intuition of the calculation is that, since all interactions take the Ising form, we can simply add up the phases. It turns out that the resulting expression, after being written in an exponential form, can be factorized into a product of cosines corresponding to the individual phase accumulations, which can then be integrated over different spatial configurations. We will not reproduce the full details here, but simply note that the same argument should also apply to a spin located at some distance external to the spin bath, and one simply needs to replace the interaction strength by the corresponding one for the external spin.

Thus, we need to calculate the interaction strength between the NV center located at $(0,0,-d_{\mathrm{NV}})$ and a spin located at position $(r\cos\alpha, r\sin\alpha, 0)$. The quantization axis is assumed to be pointing in $\hat{n} = (\sin\theta, 0, \cos\theta)$, where $\theta = 54.7^\circ$, such that $\cos^2\theta = 1/3$. Taking the inner product, we find that the interaction strength is

$$\frac{\mu_0}{4\pi}\frac{\gamma_e^2\hbar t}{2(r^2+d_{\mathrm{NV}}^2)^{3/2}}\left(\frac{3(r\cos\alpha\sin\theta + d_{\mathrm{NV}}\cos\theta)^2}{r^2+d_{\mathrm{NV}}^2} - 1\right) = \frac{\mu_0}{4\pi}\frac{\gamma_e^2\hbar t(2r^2\cos^2\alpha - r^2 - 2\sqrt{2}rd_{\mathrm{NV}}\cos\alpha)}{2(r^2+d_{\mathrm{NV}}^2)^{5/2}}. \tag{C16}$$

Transcribing the expressions found in Ref. [27] with these interactions, we find that the coherence decay should satisfy

$$\langle S(t)\rangle_c = \exp\left\{\sigma\int_0^{2\pi} d\alpha\int_0^{\infty} rdr\times\left[\cos\frac{\mu_0\gamma_e^2\hbar\tau\,(2r^2\cos^2\alpha - r^2 - 2\sqrt{2}rd_{\mathrm{NV}}\cos\alpha)}{8\pi(r^2+d_{\mathrm{NV}}^2)^{5/2}} - 1\right]\right\}$$
$$= \exp\left\{\sigma d_{\mathrm{NV}}^2\int_0^{2\pi} d\alpha\int_0^{\infty} xdx\times\left[\cos\frac{\mu_0\gamma_e^2\hbar\tau}{8\pi d_{\mathrm{NV}}^3}\frac{(2x^2\cos^2\alpha - x^2 - 2\sqrt{2}x\cos\alpha)}{(x^2+1)^{5/2}} - 1\right]\right\}, \tag{C17}$$

where we have written $x = r/d_{\mathrm{NV}}$.

We would like to note that here $\mu_0\gamma_e^2\hbar\tau/8\pi d_{\mathrm{NV}}^3$ always appears together in the integral, which means that the shape of the curve should be the same regardless of the depth d_{NV}. Changing d only leads to a rescaling of the time at which changes in shape happen. However, note that, since there is a d_{NV}^2 in front, the timescale involved in the stretched exponential decay can change.

While the above expression does not appear to be easily integrated, we can take appropriate limits to understand its behavior. First, let us consider the limit where τ is small, such that $\mu_0\gamma_e^2\hbar\tau \ll 8\pi d_{\mathrm{NV}}^3$. In this case, intuitively, the spins far away will not have the time to interact yet, so we are dominated by contributions from small r. In this case, the argument of the cosine in Eq. (C17) is small, such that we can perform a Taylor expansion to obtain

$$\langle S(t)\rangle_c = \exp\left\{\sigma d_{\mathrm{NV}}^2\int_0^{2\pi} d\alpha\int_0^{\infty} xdx\times\left[-\frac{1}{2}\left(\frac{\mu_0\gamma_e^2\hbar\tau}{8\pi d_{\mathrm{NV}}^3}\frac{(2x^2\cos^2\alpha - x^2 - 2\sqrt{2}x\cos\alpha)}{(x^2+1)^{5/2}}\right)^2\right]\right\}$$
$$= \exp\left[-\frac{3\pi\sigma d_{\mathrm{NV}}^2}{16}\left(\frac{\mu_0\gamma_e^2\hbar\tau}{8\pi d_{\mathrm{NV}}^3}\right)^2\right]. \tag{C18}$$

Thus, at short times, we recover the Gaussian decay in Eq. (C15).

Next, let us consider the limit where τ is large, such that $\mu_0\gamma_e^2\hbar\tau \gg 8\pi d_{\mathrm{NV}}^3$. Intuitively, the time is long enough such that the external spin sees the influence of all spins far away, and the situation will be similar to a 2D spin bath. In this case, we have contributions from a large range of values of x. Because of the x factor outside the parentheses however, coming from the volume element in the Jacobian, the main contribution to the integral will be coming from values of x that are large, and thus we may approximate the fraction as

$$\frac{2x^2\cos^2\alpha - x^2 - 2\sqrt{2}x\cos\alpha}{(x^2+1)^{5/2}} \approx \frac{2x^2\cos^2\alpha - x^2}{x^5} = \frac{2\cos^2\alpha - 1}{x^3}. \tag{C19}$$

Explicitly performing the integral in this limit gives

$$\langle S(t)\rangle_c = \exp\left(-\frac{9\sqrt{\pi}\Gamma(11/6)\sigma}{5}\left(\frac{\mu_0\gamma_e^2\hbar\tau}{8\pi}\right)^{2/3}\right), \tag{C20}$$

which reproduces the $\tau^{2/3}$ scaling in this limit. Note also that in this limit, d_{NV} drops out of the problem. The intuition is again that the spin sees everything out to infinity, and the only spatial scale left in the problem is an average distance between the surface spins.

APPENDIX D: ADDITIONAL SURFACE SPIN MEASUREMENTS

1. Coherently coupled surface spins

Over all the samples investigated here, coherently coupled surface spins are rarely observed. One example of this rare occurrence is shown in the main text, and we present data from a second NV center in Fig. 9. Based on the coupling strength, it is 6–7 nm from the NV center, which is consistent with the 7 nm depth as measured by proton NMR. However, we also note that it cannot be conclusively shown with our methods that this defect exists at the surface. This particular spin survived multiple acid cleanings and surface treatments and so may be deeper than the bath spins probed in the remainder of this work. There may exist several types of defects with an observable DEER signal, some of which are found below the diamond surface [5]. It may be that these slightly deeper defects are also more stable.

2. Surface spin linewidth

Another important parameter of the surface spin bath is the linewidth of the DEER transition, as this provides information about the disorder on the surface experienced

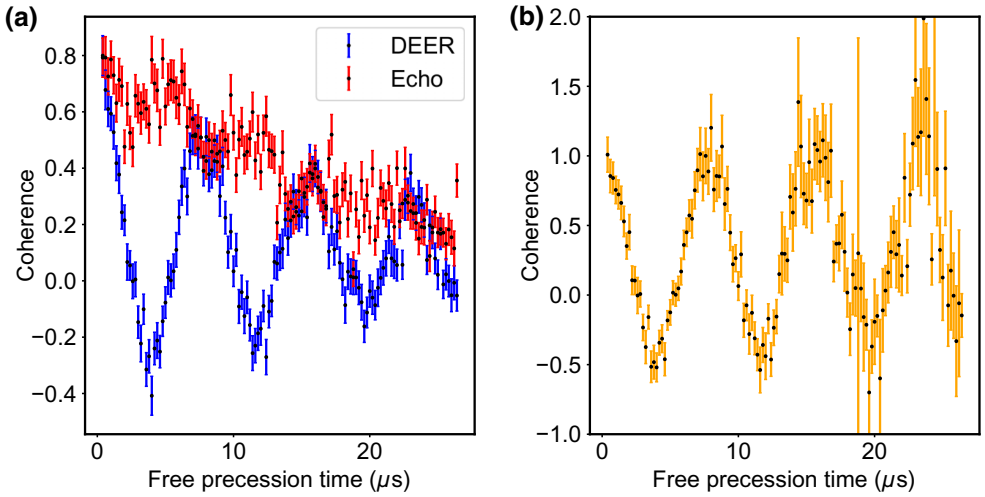


FIG. 9. Coherently coupled surface spin. (a) Raw DEER (blue) and echo (red) data from a NV center that is coherently coupled to a surface spin. (b) The raw DEER signal divided by the echo signal.

by the surface spins. We measure the linewidth by performing a DEER pulse sequence while sweeping the frequency of the surface spin π pulse. A typical dataset, with a Lorentzian fit, is shown in Fig. 10. We note that this value of 35 MHz is slightly larger than the approximate 20 MHz linewidths reported in the literature [5,7,11,12], which may be a result of power broadening. For the following discussions, this discrepancy will not be relevant, however.

3. Surface spin $T_{2,\mathrm{Rabi}}$

As an additional measurement to bound the disorder of the surface spins, we sweep the duration of the microwave pulse performed on the surface spins during the DEER sequence. We fit the resulting curve to a cosine with an exponentially decaying envelope, and find a characteristic decay time of 200 ± 30 ns (see Fig. 11).

4. Measured surface spin T_1 relaxation times

The surface spin T_1 relaxation times are measured on several NV centers in Sample 9 using a correlation spin sequence [7,56]. Results can be seen in Fig. 12. Additionally, for Sample 3 and Sample 8, surface spin T_1 times are measured by sweeping the time of the DEER microwave pulse within the echo sequence [11]. Extracted T_1 times are consistent with those measured in correlation spectroscopy (see Fig. 13). These T_1 times are significantly longer than the length of the DEER sequences performed in the main text.

We also note that these T_1 times place a lower bound on the in-sequence reconfiguration time. These values are orders of magnitude longer than the correlation times of the traps that have been identified in electric field noise measurements of shallow NV centers [57].

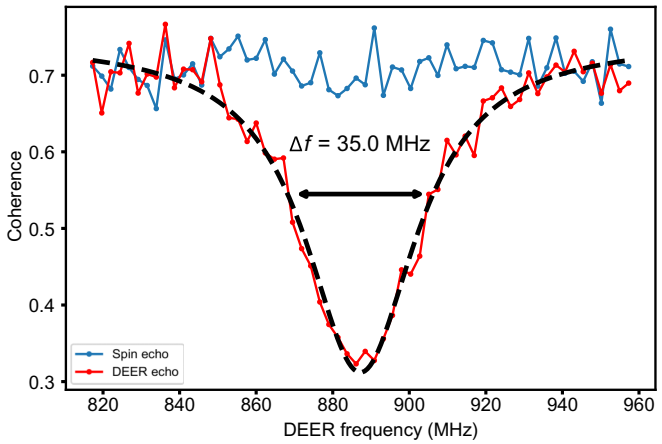


FIG. 10. DEER frequency sweep.

APPENDIX E: MEAN NEAREST-NEIGHBOR SEPARATION

In the main text, an average spin-spin separation, ℓ_{SS}, is calculated from a fitted density of surface spins, σ. From dimensional analysis, these quantities are related by

$$\ell_{\mathrm{SS}} \propto \frac{1}{\sqrt{\sigma}}; \tag{E1}$$

however, the constant of proportionality is ambiguous. In the main text, we specifically refer to the mean value of the distribution of nearest-neighbor distances for a randomly distributed layer of spins with density σ. In this case the constant of proportionality is $1/2$, and for completeness, we derive this result below.

We can obtain this by considering the probability of finding a neighboring spin a distance r away. The distribution function $w(r)dr = P(\text{no neighbor} < r) \times P(\text{neighbor between } r \text{ and } r + dr)$:

$$w(r)dr = \left(1 - \int_0^r w(r')dr'\right) \times (2\pi r dr \sigma). \tag{E2}$$

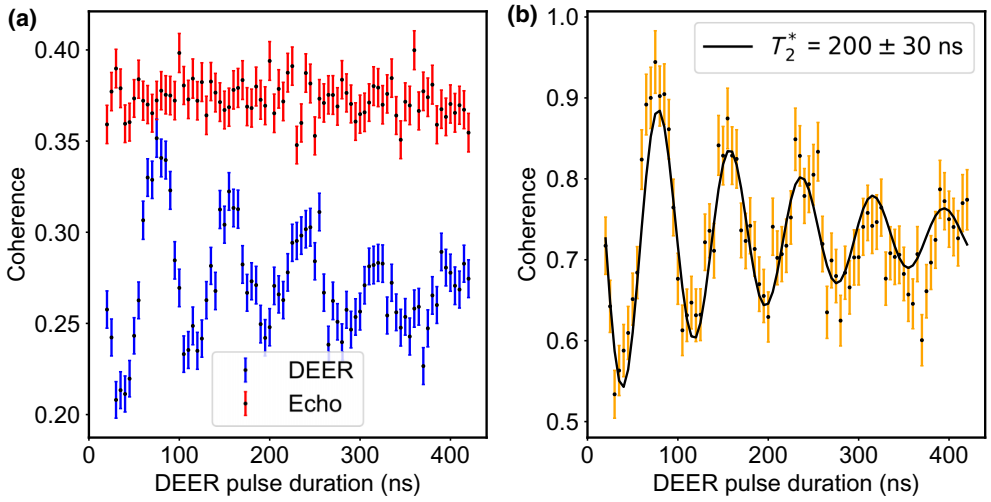


FIG. 11. DEER Rabi measurement. (a) We perform a microwave pulse of varying length on the surface spins during the DEER sequence and (b) fit the normalized signal to a decaying exponential envelope to extract a coherence decay time.

Taking the derivative gives

$$\frac{dw(r)}{dr} = 2\pi\sigma\left(1 - \int_0^r w(r')dr'\right) - w(r) \times (2\pi r\sigma). \tag{E3}$$

Substituting into the first equation we can eliminate the integral

$$\frac{dw(r)}{dr} = \frac{w(r)}{r} - w(r) \times (2\pi r\sigma). \tag{E4}$$

This differential equation has the solution

$$w(r) = Kr\exp\{-\pi r^2\sigma\}, \tag{E5}$$

where K is an integration constant, which from normalization must be $K = 2\pi\sigma$. So the nearest-neighbor probability distribution function is

$$w(r) = 2\pi\sigma r\exp\{-\pi r^2\sigma\}. \tag{E6}$$

We then calculate the first moment of Eq. (E6):

$$\begin{aligned}\langle \ell_{SS}\rangle &= \int_0^\infty rw(r)dr \\ &= 2\pi\sigma\frac{1}{4\pi\sigma^{3/2}} \\ &= \frac{1}{2\sqrt{\sigma}}.\end{aligned} \tag{E7}$$

Equation (E7) is the expression used in the calculations of the main text.

APPENDIX F: EFFECTS OF FINITE SURFACE SPIN T_1 TIME

Here we calculate the effects of finite surface spin T_1 time on the DEER exponent n in order to investigate

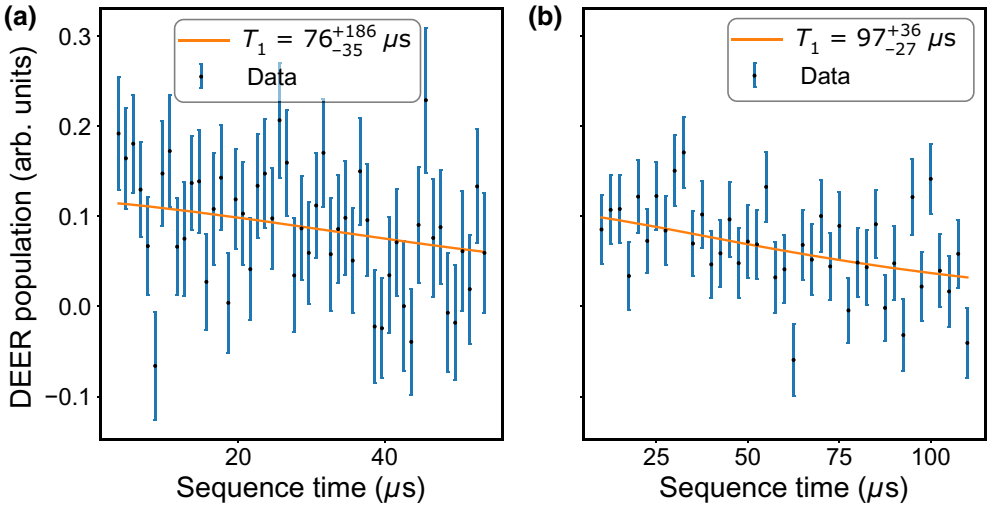


FIG. 12. Measured T_1 of surface spins using a correlation sequence. Relaxation time T_1 for surface spins measured by two different NV centers, depicted in (a) and (b), using the correlation sequence described in Ref. [7].

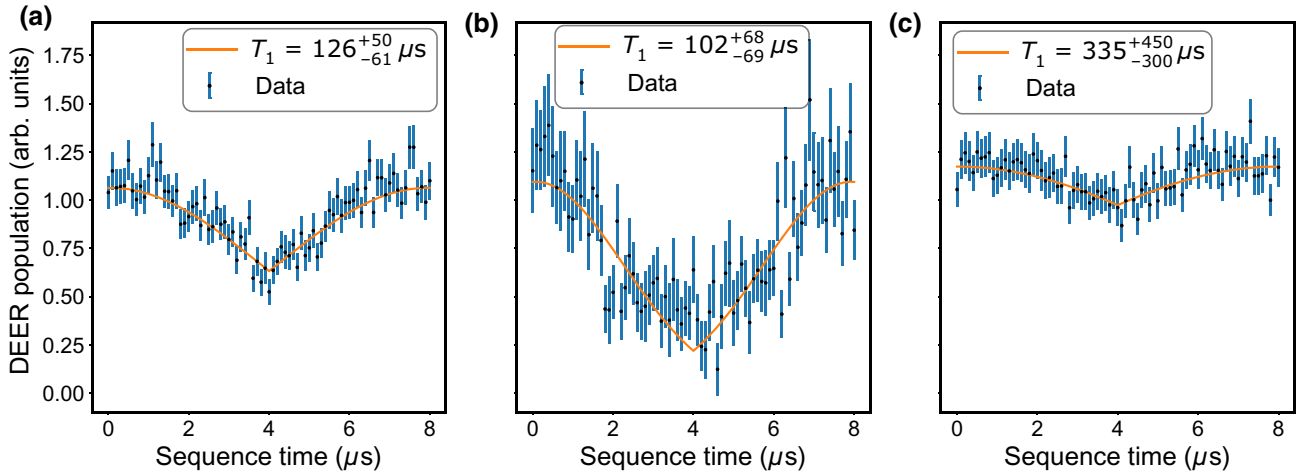


FIG. 13. Measured T_1 of surface spins by a bath correlation measurement. Relaxation time T_1 for surface spins measured by three different NV centers, depicted in (a)–(c), using the sequence described in Ref. [11]. Panels (a) and (b) are from Sample 8 and (c) is from Sample 3.

whether the observed coherence decay behavior can be accounted for by surface spin relaxation.

In the presence of Gaussian-distributed noise amplitudes (as expected from a continuous density of surface spins), the quasistatic limit $T_1 \gg T_{2,\mathrm{DEER}}$ gives a DEER decay exponent $n = 2$, and the Markovian limit $T_1 \ll T_{2,\mathrm{DEER}}$ gives a DEER decay exponent $n = 1$. Quantitatively, the case of finite-T_1 time is calculated in Ref. [11] and is given by

$$\langle S(\tau)\rangle = \exp\left[-(\Gamma_{\mathrm{DEER}}\tau)^2 f(\tau, T_1)\right], \tag{F1}$$

where

$$f(\tau, T_1) = \frac{2T_1^2}{\tau^2}\left[\frac{\tau}{T_1} - 1 + e^{-\tau/T_1}\right] \tag{F2}$$

is a dimensionless quantity that describes the effects of finite surface spin T_1 relaxation time. Note that in the limit of $\tau \ll T_1$ one finds $f(\tau, T_1) = 1$.

In Fig. 14 we fit Eq. (F1) to $\exp(-(\tau/T_{2,\mathrm{DEER}})^n)$. As expected, for shallow NV centers $T_{2,\mathrm{DEER}} \ll T_1$ and so we recover the quasistatic limit $n = 2$, whereas for deeper NV centers $T_{2,\mathrm{DEER}} \gg T_1$ and we recover the Markovian limit $n = 1$. The trend of n vs depth in Fig. 14(b) is opposite

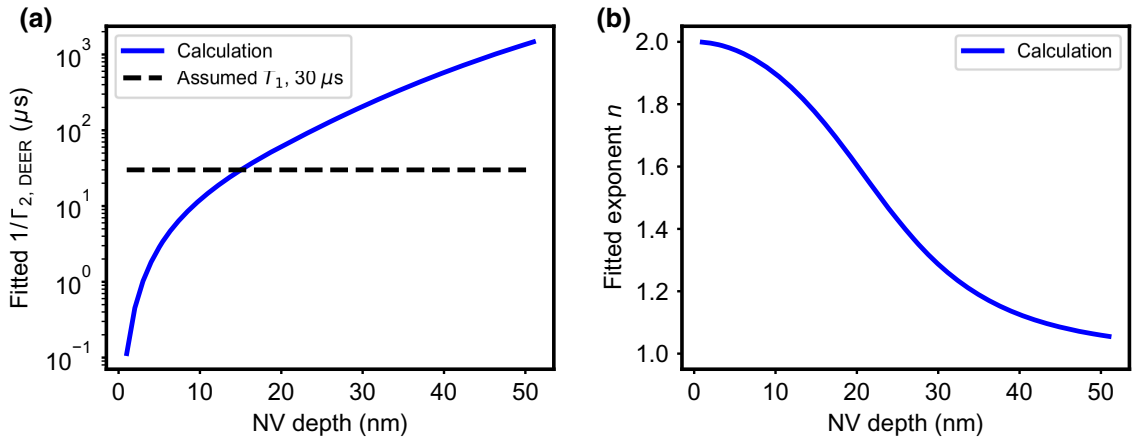


FIG. 14. Depth scaling in the presence of finite surface spin T_1 time. (a) Calculated decay timescale $T_{2,\mathrm{DEER}}$ with decay profile calculated using Eq. (F1) and fit to $\exp(-(\tau/T_{2,\mathrm{DEER}})^n)$, assuming a finite surface spin T_1 time of 30 μs (black dashed line) and a density 0.005 nm^{-2}. (b) Fitted exponent n. Once $T_{2,\mathrm{DEER}}$ becomes longer than the surface spin T_1 relaxation time, n decreases from the infinite-T_1 limit of $n = 2$ toward the Markovian limit of $n = 1$. The calculated increase of n for shallower NV centers is opposite of the observed trend versus depth in Fig. 2(c) of the main text, and suggests that the depth trend in Fig. 2(c) is not explained by the effects of finite relaxation time.

from the trend observed in Fig. 2(c) of the main text, indicating that spin relaxation cannot account for our data.

1. Exponents *n* under the assumption of a noise spectral density

Here we show that the exponents n from DEER exponential decays of the form $C(\tau) = \exp[-(\tau/T_2)^n]$ will have $n \geq 1$ if the noise spectrum is flat ($n = 1$) or monotonically decreasing as a function of frequency ($n > 1$), assuming the noise spectral density formulation holds. The assumption of a monotonically decreasing noise spectrum is physically reasonable and consistent with experimentally observed noise spectra; accordingly, a likely explanation for the observed DEER exponents $n < 1$ is that the noise spectral density formulation does not hold, in particular because the noise amplitudes are not Gaussian distributed, as is found in the case of discrete hopping spins.

We now show that $n \geq 1$ for a monotonically decreasing noise spectral density. We show this to be true for a general N-pulse qubit filter function $F_N(\omega T)/\omega^2$, defined to be continuous and non-negative across all frequencies, including Ramsey and DEER measurements with filter function $4\sin^2(\omega T/2)/\omega^2$. Note that during a DEER sequence, if the surface spins are perfectly flipped during the NV center π pulse then the NV center filter function to the surface spins will be identical to that of Ramsey. The FID can then be described by Eq. (B1). The noise spectrum $S(\omega) \geq 0$, and for physically relevant noise sources is monotonically decreasing, i.e., $dS(\omega)/d\omega \leq 0$ for all ω, and with $S(\omega)$ and $dS(\omega)/d\omega$ defined over $\omega \in [0, \infty)$. We rewrite Eq. (B2) as

$$\chi(T) = \frac{1}{\pi} T\left[\int_0^\infty S\left(\frac{x}{T}\right)\frac{F_N(x)}{x^2}dx\right], \tag{F3}$$

where $x \equiv \omega T$ is a dimensionless integration variable. In order to probe the exponent n associated with $\chi(T) = (T/T_2)^n$, we apply $(Td/dT - 1)$ to $\chi(T)$ in Eq. (F3); using the product rule, we find that

$$\left(T\frac{d}{dT} - 1\right)\chi(T) = \frac{1}{\pi}T^2\left[\int_0^\infty \frac{dS(x/T)}{dT}\frac{F_N(x)}{x^2}dx\right]. \tag{F4}$$

Using $\omega(T) = x/T$, we can evaluate

$$\frac{dS[\omega(T)]}{dT} = \frac{dS(\omega)}{d\omega}\frac{d\omega(T)}{dT} = \frac{dS(\omega)}{d\omega}\left(\frac{-x}{T^2}\right). \tag{F5}$$

Equation (F4) simplifies to

$$T\frac{d\chi(T)}{dT} - \chi(T) = \frac{1}{\pi}\int_0^\infty \left(\frac{-dS(\omega)}{d\omega}\right)\Bigg|_{\omega=x/T}\frac{F_N(x)}{x}dx. \tag{F6}$$

We can rewrite the left-hand side of Eq. (F6) by using $dT/T = d\ln T$, and since, by assumption, $\chi(T)$ is positive and differentiable at times $T > 0$, we write

$$\begin{aligned} T\frac{d\chi(T)}{dT} - \chi(T) &= \chi(T)\left(\frac{1}{\chi(T)}\frac{d\chi(T)}{d\ln T} - 1\right) \\ &= \chi(T)\left(\frac{d\ln\chi(T)}{d\ln T} - 1\right). \end{aligned} \tag{F7}$$

So, Eq. (F6) can be written as

$$\frac{d\ln\chi(T)}{d\ln T} = 1 + \frac{1}{\chi(T)}\frac{1}{\pi}\int_0^\infty \left(\frac{-dS(\omega)}{d\omega}\right)\Bigg|_{\omega=x/T}\frac{F_N(x)}{x}dx. \tag{F8}$$

As defined, the right-hand side is necessarily ≥ 1 for all times T because $\chi(T) > 0$ for all T, $F_N(x) \geq 0$ for all x, and $dS(\omega)/d\omega \leq 0$ for all ω. Accordingly,

$$\frac{d\ln\chi(T)}{d\ln T} \geq 1 \tag{F9}$$

for all free precession times $T > 0$. And, if in some region of times T the function $\chi(T)$ has a power-law form $(T/T_2)^n$ then

$$n \geq 1. \tag{F10}$$

Under these mathematical arguments, we cannot observe $n < 1$ for a monotonically decreasing noise spectral density, because the integral is shown to be ≥ 0 for all times T. A monotonically increasing noise spectral density would give a negative integral in Eq. (F8) and therefore $n < 1$, but this is physically unreasonable and inconsistent with observed noise spectra.

Furthermore, note that Eq. (F9) holds for a general $\chi(T)$, and implies that the slope on a plot of $\ln\chi(T)$ versus $\ln(T)$ will be ≥ 1 for a monotonically decreasing noise spectral density. In the main text we plot $\ln\chi(T)$ versus $\ln(T)$ and find slopes less than 1, which implies that the noise spectral density formulation does not hold, in particular because the noise amplitudes are not Gaussian distributed. The configurational averaging of surface spins results in a non-Gaussian distribution of noise amplitudes (i.e., for some shots, there are many surface spins nearby, and for some shots, there are no surface spins nearby), which violates the requirements for the noise spectral density formulation and explains the exponents $n < 1$ that we observe.

2. Note about filter functions

In the above derivation, it is assumed that the filter function used during a measurement is that of the Ramsey FID due to the surface spins alone. Experimentally, this is not what is measured due to nonsurface spin contributions to

the FID. In the main text, we attempt to remedy this by dividing by a spin-echo signal. However, depending on the form of the noise spectrum, this filter function may be different from what is desired.

To see this, we write the noise spectrum experienced by the NV center as a sum of two terms:

$$S(\omega) = S(\omega)_{\mathrm{SS}} + S(\omega)_{\mathrm{EE}} \tag{F11}$$

with $S(\omega)_{\mathrm{SS}}$ the noise spectrum of the surface spins and $S(\omega)_{\mathrm{EE}}$ the noise spectrum due to all other noise sources. The spin-echo measurement probes the entire noise spectrum $S(\omega)$ with the spin-echo filter function. In the spectral noise formalism the signal is given by

$$C(\tau) = \exp\frac{-1}{\pi}\int_0^\infty S(\omega)F^{\mathrm{SE}}(\omega\tau)\frac{d\omega}{\omega^2}, \tag{F12}$$

where $F^{\mathrm{SE}}(\omega\tau)$ is the spin-echo filter function, given by

$$F^{\mathrm{SE}}(\omega\tau) = 8\sin^4\frac{\omega\tau}{4}. \tag{F13}$$

The DEER measurement, in contrast, probes $S(\omega)_{\mathrm{SS}}$ with the spin-echo filter function, but $S(\omega)_{\mathrm{EE}}$ is probed with the Ramsey filter function, $F^R(\omega)$, given by

$$F^R(\omega\tau) = 2\sin^2\frac{\omega\tau}{2}. \tag{F14}$$

The DEER measurement signal is then given by

$$D(\tau) = \exp\frac{-1}{\pi}\int_0^\infty (S(\omega)_{\mathrm{SS}}F^R(\omega\tau) + S(\omega)_{\mathrm{EE}}F^{\mathrm{SE}}(\omega\tau))\frac{d\omega}{\omega^2}. \tag{F15}$$

When we divide these two signals, what we actually measure is not the FID of the surface spins alone, but rather the quantity

$$S(\tau) = \exp\frac{-1}{\pi}\int_0^\infty S(\omega)_{\mathrm{SS}}(F^R(\omega\tau) - F^{\mathrm{SE}}(\omega\tau))\frac{d\omega}{\omega^2}. \tag{F16}$$

If the surface spins have a long relaxation time ($T_1^{\mathrm{DEER}} > T_2^{\mathrm{NV}}$), the second term is negligible due to the low weight of the filter function near $\omega = 0$. For fast relaxation times, the stretching factor of the quantity $S(\tau)$ does not go to 2/3, but rather to 0, and this occurs for $\tau > T_1^{\mathrm{DEER}}$. This can be intuitively understood by the following argument: once the free precession interval exceeds T_1^{DEER}, there is no difference between the measurement with and without a π pulse on the surface spins and so the two measurements become nominally identical. Our measurements of surface spin relaxation times (see Appendix D 4) indicate that we are not in this regime, but for systems with much shorter relaxation times, this additional transition should be observable.

3. Effects of interactions

Spin flips among surface spins can arise from environmental interactions and interactions between the surface spins (e.g., dipolar flip flops). In order to determine whether flip flops among static surface spins could reproduce the observed experimental results, we perform Monte Carlo simulations for abut 100 random fixed spin configurations. Interactions between the spins are assumed to be incoherent due to the large mismatch between the measured disorder of surface spin systems [$1/(2\pi T_2^*) \approx$ 35 MHz; see Appendix D 2] and the average nearest-neighbor dipolar couplings at the inferred spin densities (approximately 100 kHz). In this regime, the probability

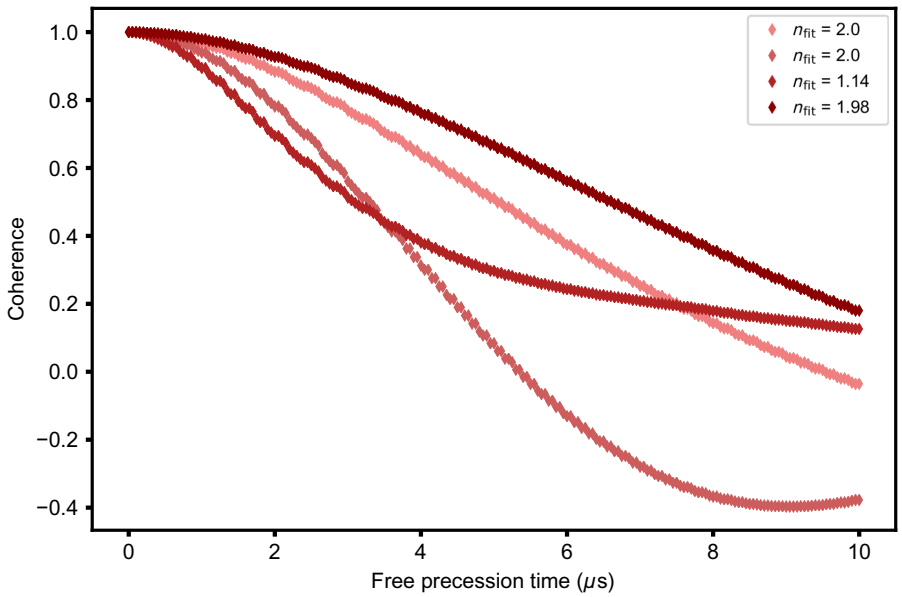
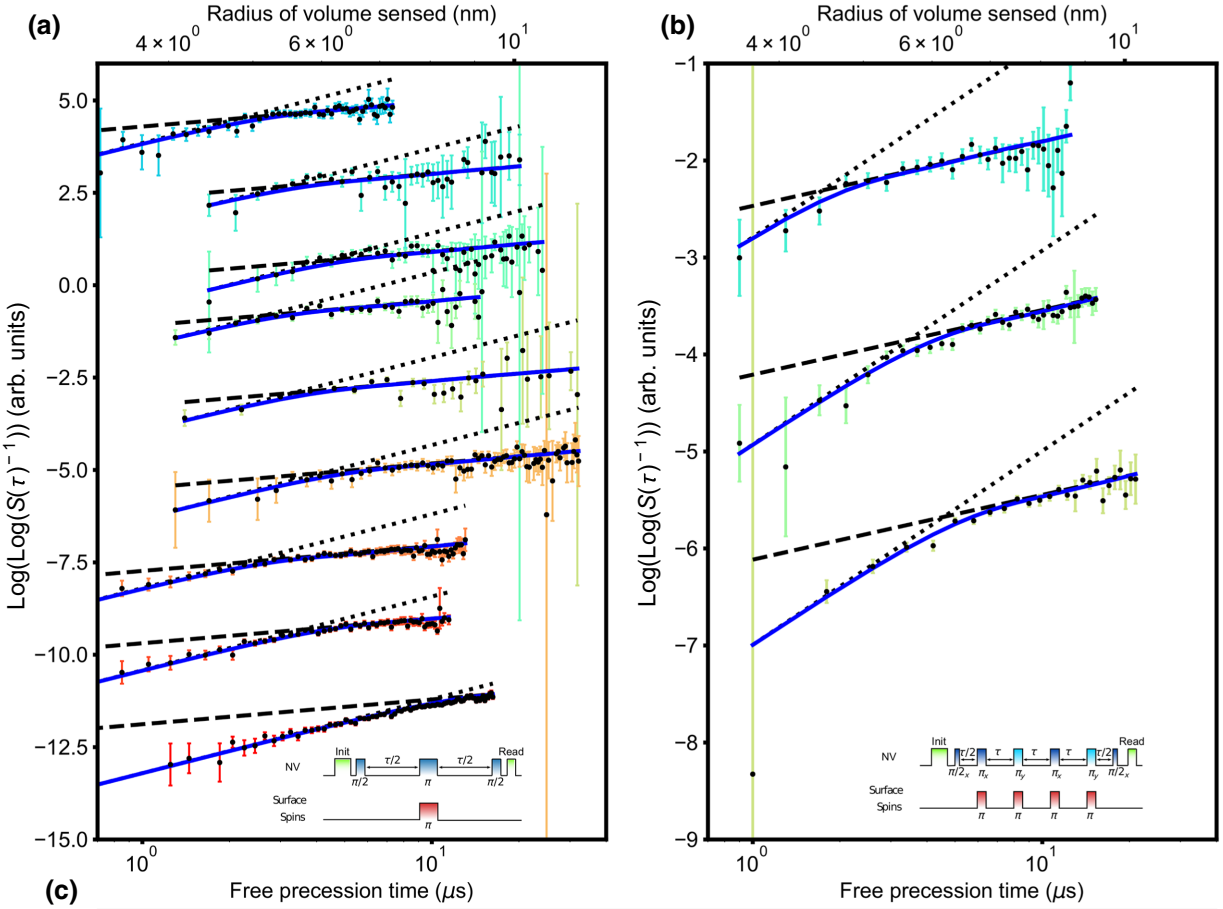


FIG. 15. Simulations with incoherent interactions. Example DEER curves for simulations allowing both flip flops and spin lattice relaxation between surface spins. Each color represents a different randomly sampled surface spin configuration. Fitting these curves with a stretched exponential produces values of $1 < n < 2$. Some curves additionally show coherent oscillations if the coupling rate to the nearest spin exceeds the rate of incoherent decay.

of a spin exchange between spin i and j is given by [58]

$$p_{i,j} = \frac{\omega_{i,j}^2}{\gamma}\delta t, \tag{F17}$$

where $\omega_{i,j}$ is the dipolar coupling between spins i and j and γ is the disorder. Here δt is the time step of the simulation, chosen such that all $p_{i,j} < 1$. Additionally, spins are allowed to relax without spin exchange through spin-lattice relaxation at a rate $1/(2T_1)$. For a simulation, spins are given random coordinates, the all-to-all couplings are calculated, and the spins randomly initialized. At each time step, spins are allowed to flip or flip flop based on draws from a Bernoulli distribution, and the phases for the entire time evolution are summed. Some results for simulation parameters $d_{\mathrm{NV}} = 5$ nm, $\sigma_{2D} = 0.005$ nm^{-2}, $T_1 = 50$ μs,



(c)

	sample	depth $_{\mathrm{NMR}}$ (nm)	depth $_{\mathrm{DEEREcho}}$ (nm)	density $_{\mathrm{DEEREcho}}$ (nm^{-2})	depth $_{\mathrm{DEERXY4}}$ (nm)	density $_{\mathrm{DEERXY4}}$ (nm^{-2})
1	9	$3.6^{+0.3}_{-0.3}$	$4.7^{+0.3}_{-0.3}$	$0.005^{+0.0009}_{-0.0006}$	...	...
2	3	$3.6^{+0.1}_{-0.1}$	$5.1^{+0.5}_{-0.8}$	$0.001^{+0.0002}_{-0.0001}$	$4.3^{+0.6}_{-1.4}$	$0.001^{+0.0002}_{-0.0001}$
3	3	$4.0^{+0.2}_{-0.2}$	$5.8^{+11.7}_{-0.8}$	$0.000^{+0.0083}_{-0.0001}$	...	...
4	3	$4.2^{+0.2}_{-0.2}$	$4.9^{+0.4}_{-0.3}$	$0.002^{+0.0003}_{-0.0002}$	$5.5^{+0.5}_{-0.6}$	$0.002^{+0.0002}_{-0.0002}$
5	3	$4.4^{+0.2}_{-0.2}$	$5.4^{+7.5}_{-0.6}$	$0.001^{+0.0211}_{-0.0002}$	$6.0^{+0.4}_{-0.6}$	$0.002^{+0.0004}_{-0.0002}$
6	3	$4.9^{+0.2}_{-0.2}$	$5.8^{+9.2}_{-0.7}$	$0.001^{+0.0071}_{-0.0001}$	...	...
7	9	$5.3^{+1.2}_{-1.2}$	$4.7^{+0.3}_{-0.3}$	$0.005^{+0.0009}_{-0.0005}$	...	...
8	9	$6.5^{+0.3}_{-0.3}$	$5.6^{+0.4}_{-0.3}$	$0.005^{+0.0013}_{-0.0007}$	...	...
9	9	$9.3^{+0.6}_{-0.6}$	$7.9^{+0.6}_{-0.4}$	$0.004^{+0.0012}_{-0.0007}$	...	...

FIG. 16. Additional NV center measurements. (a) Plots of $\log S(\tau)^{-1}$ for nine additional NV centers. Inset: pulse sequence used to measure NV centers in (a). Blue lines are fits to the hopping model in the main text, and dotted and dashed lines show slopes of 2 and 2/3, respectively. The y-axis offset is arbitrary. (b) DEER XY-4 data for three of the nine additional NV centers, plotted as in (a). Inset: DEER XY-4 pulse sequence. (c) Surface spin densities and depths extracted from NV centers in (a) and (b). Here $\mathrm{Depth}_{\mathrm{NMR}}$ is the NV center depth extracted from fitting the proton NMR signal, while $\mathrm{Depth}_{\mathrm{DEER}}$ and $\mathrm{Depth}_{\mathrm{DEERXY4}}$ are the depths extracted from the hopping model in the main text. Error bars are one standard error.

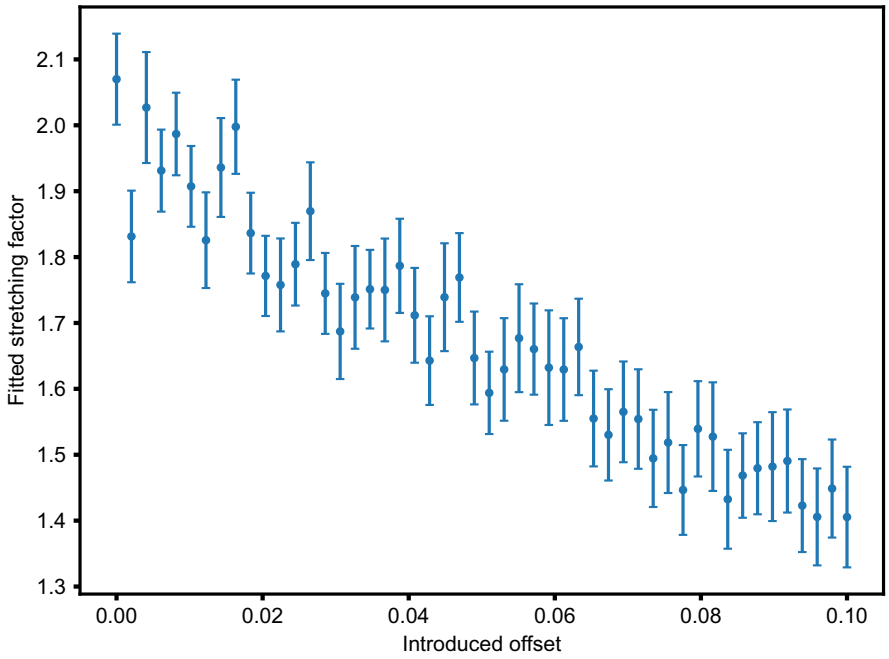


FIG. 17. Lowering of stretching factors with introduced offset. In simulated data with an introduced offset, the fitted values of n underestimate the true value of $n = 2$ systematically, with lower values of n fitted for larger introduced offsets. The data are fit assuming an offset $d = 0$.

and $\gamma = 100$ kHz are shown in Fig. 15. For these parameters, about 20% of configurations lead to the observation of coherent oscillations in the DEER signal (whenever the nearest spin coupling is greater than the incoherent decay rate). No simulations with experimentally relevant parameters have resulted in stretching factor $n < 1$.

APPENDIX G: EXPONENT TRANSITION FOR ADDITIONAL NV CENTERS

In Fig. 16, we present additional finely sampled datasets from NV centers of various depths as in Fig. 3(a). Again, we plot $\log S(\tau)^{-1}$ on a log-log scale so that the exponent n is given by the slope of the line. In addition to the DEER echo-type measurements presented in the main text, for some NV centers, we perform a dynamical decoupling sequence (XY-4) on the NV center while applying multiple π pulses to the surface spins, which in principle permits a longer interrogation of the surface spin FID before the NV center decoheres. The interpretation of this type of measurement is complicated by the short dephasing time of the surface spins and the longer total rotation time required for multiple π pulses. The close agreement with $n = 2/3$ at long times is striking nevertheless.

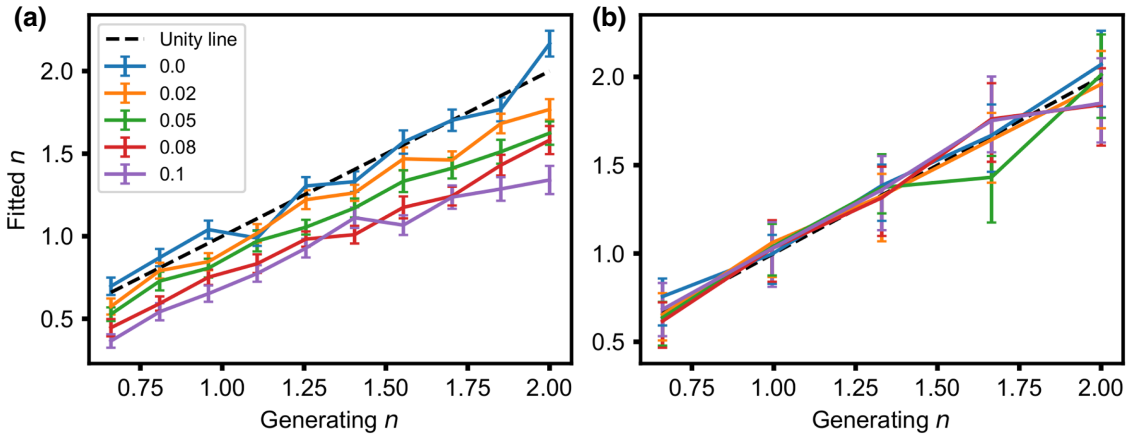


FIG. 18. Recovery of actual stretching factors. (a) When an offset term is included in the fit, the true stretching factor is not fully recovered when fitting the DEER echo divided by the coherence when the offset is large. (b) Simultaneously fitting the coherence and DEER signals using MCMC, the correct stretching factor is recovered.

APPENDIX H: ADDITIONAL MODEL ANALYSIS

1. Effect of offset term

The DEER datasets in the main text are fit to a function of the form

$$y = a\exp(-(\Gamma\tau)^n) + d. \tag{H1}$$

Though we have taken great care to avoid introducing an offset during data acquisition and processing, we fit to $d \neq 0$ because an offset term can artificially lower the fitted value of the stretching factor. To explore this phenomenon, we generate simulated data with a known value of n, where the coherence and DEER signals are both of the form of Eq. (H1). This is used to sample from two Poisson distributions with different parameters, and the resulting simulation is processed and fit as if it were from an actual NV center. In particular, the simulated DEER echo signal is divided point by point by the simulated coherence to obtain a DEER FID curve. We find in Fig. 17 that if d is constrained to be zero, but an artificial offset does exist, then the fitted values of n are lower than that used to generate the simulated data ($n = 2$).

The introduction of an offset term partially solves this problem, as is shown in Fig. 18(a). There remains some dependence of the fitted stretching factor on the introduced offset value, however, and again the value of n is lower than that used to generate the data for positive values of the offset. By fitting the coherence and DEER echo signals to separate likelihood functions using Markov chain Monte Carlo, where the DEER signal is assumed to result from a DEER FID multiplied by the coherence, this offset dependence vanishes and the true generating n is recovered [Fig. 18(b)].

In Fig. 19, we compare these fitting methods on a subset of actual (not simulated) data from isotpically purified diamonds, where the simultaneous fits are possible due to the absence of ^{13}C induced coherence oscillations. We find that the introduction of an offset term to our actual data has little effect on the fitted value of the stretching factor, which implies that the offsets are negligible in these samples. Including the offset term does introduce larger errors, however. Fitting the coherence and DEER echo together using the double likelihood MCMC method produces results with smaller errors, but it systematically fits a slightly higher value of n, as might be expected from the above analysis. We have therefore chosen to fit as many datasets as possible using this method, with the remainder fit after division due to oscillations in the coherence that are not captured by the simple exponential model. We note that the observed trends in stretching factors remain, despite the majority of the data being fit by a model that systematically fits a higher value of n.

2. Analysis of two-exponent fits

For the shallow NV centers investigated in this work, a single exponential with $n = 2$ fixed leads to a poor fit while permitting n to vary leads to better fits with $n < 2$, in general. Additionally, we claim that, for configurationally averaged surface spins, the short and long time behaviors of the FID should have different stretching factors. In particular, it should be possible to observe $n = 2$ for short times and $n = 2/3$ for long times in NV centers that are closer to the surface than the average surface spin separation. Here we investigate the veracity and robustness of this claim by fitting the DEER signal of four NV centers from Sample 9. These are the three NV centers depicted

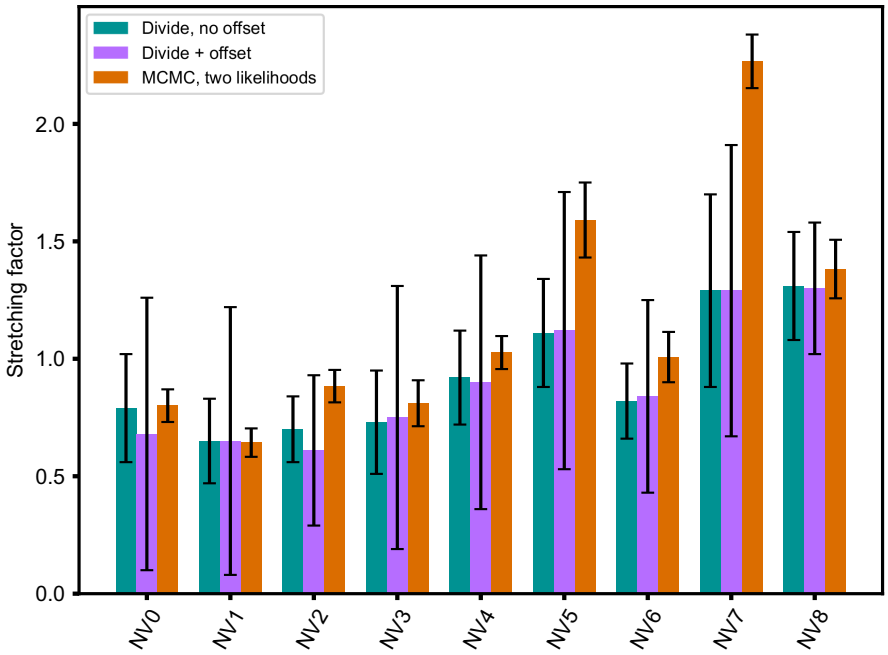


FIG. 19. Comparison of fitting methods. Sample of fitted stretching factors for NV centers in isotopically purified samples. Including an offset does not generally change the fitted value of the stretching factor, but it does introduce additional error. MCMC fitting using two likelihoods produces more constrained parameter fits, and systematically fits slightly higher values of n.

in Fig. 3 of the main text, plus one more that is located at a similar depth to the shallowest. We fit these data to a function of the form

$$f(\tau) = a\exp -(\Gamma_1\tau)^{n_1} + (1-a)\exp -(\Gamma_2\tau)^{n_2}, \quad \text{(H2)}$$

where τ is the free precession time and $a, \Gamma_1, \Gamma_2, n_1, n_2$ are the fit parameters. Here we take a fully Bayesian approach and use MCMC to sample the posterior distribution of these parameters, using a Gaussian likelihood with $\mu = f(\tau)$ and $\sigma(\tau)$ from the actual variance of the data. Priors for the fit parameters are uniform distributions. The resulting posterior distributions of n_1 and n_2 are shown in Fig. 20. We find that the early time exponent n_2 has a mean value close to two while the late time exponent n_1 is lower. For the deepest NV center, the two distributions are nearly indistinguishable, indicating that the biexponential has collapsed into a single exponential, which is expected

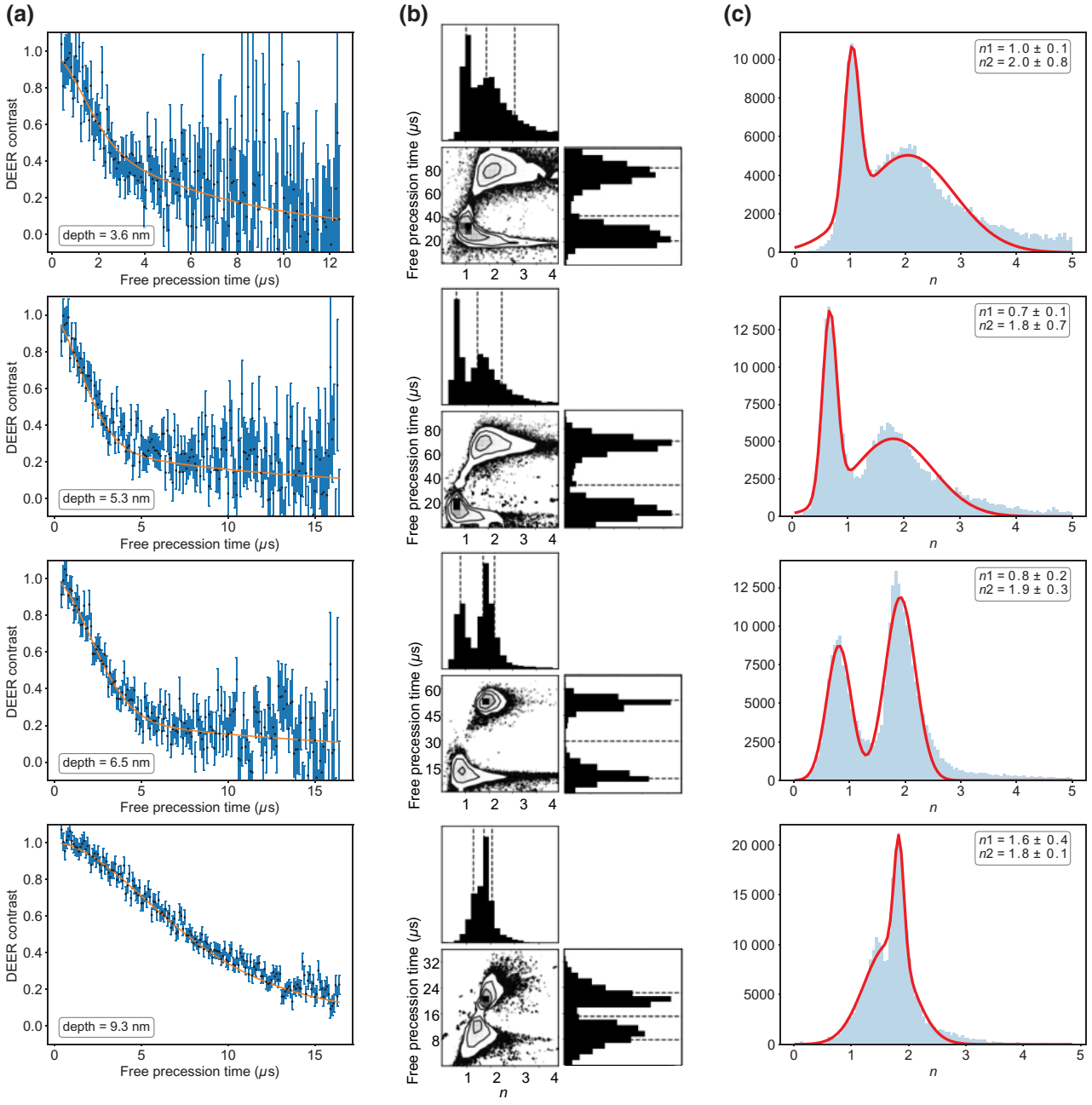


FIG. 20. Fits to a biexponential for four NV centers. (a) Data for each of the four NV centers. The orange lines are best biexponential fits based on MCMC analysis. (b) The posterior distribution function (PDF) of n versus coupling, shown for each NV. (c) The exponent stretching factors are found by fitting the PDF of n to two Gaussians.

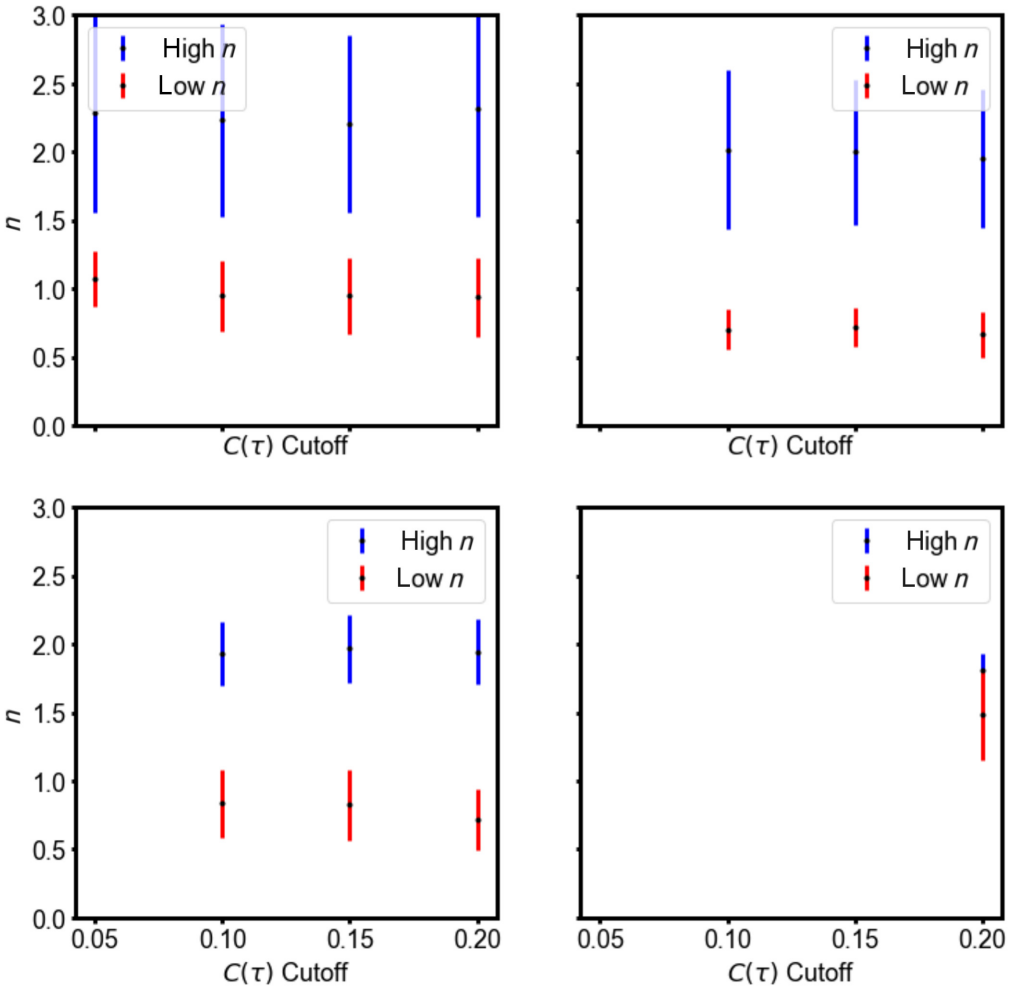


FIG. 21. Stability of stretching factors. Mean and standard error of the two stretching factors (high n and low n) of Eq. (H2) as a function of the cutoff value of the coherence for the four NV centers in Fig. 20. For two NV centers, $C(\tau)$ does not dip below 0.1 during the measurement period and for one NV center, $C(\tau)$ does not dip below 0.2 during the measurement period. However, extracted exponents remain stable for all NV centers regardless of the exact $C(\tau)$ cutoff value.

for a NV center that is deeper than the average spin-spin separation.

In processing the data to fit these stretching factors, we choose a cutoff value of $C(\tau)$ in time after which data are discarded. This is to prevent the signal-to-noise ratio of the FID data from becoming too large, which is an unavoidable consequence of dividing the DEER echo by the spin-echo coherence. We show in Fig. 21 that reasonable choices of the cutoff value do not affect the results of the fit. That is, for these datasets, there is enough coherence left at the cutoff point that the fit parameters are well constrained.

3. Model comparison

While it is instructive to extract biexponential stretching factors in order to observe short and long time behavior, Eq. (H2) is not the form that the free induction decay is

TABLE II. Model comparisons between static spins (single exponent) and configurational averaging models for four highly sampled NV center datasets.

NV center	Model	WAIC	σ_{WAIC}	LOO	σ_{LOO}	Log-likelihood	$\sigma_{\text{Log-likelihood}}$
NV1 (3.6 nm)	Hopping	196	12	196	11	1243	13
	Static	192	12	192	11	1235	14
NV2 (5.3 nm)	Hopping	253	9	253	9	1429	15
	Static	226	2	226	12	1401	17
NV3 (6.5 nm)	Hopping	277	8	277	9	1543	15
	Static	244	11	244	12	1497	19
NV4 (9.3 nm)	Hopping	293	20	293	20	1664	18
	Static	288	22	288	22	1501	72

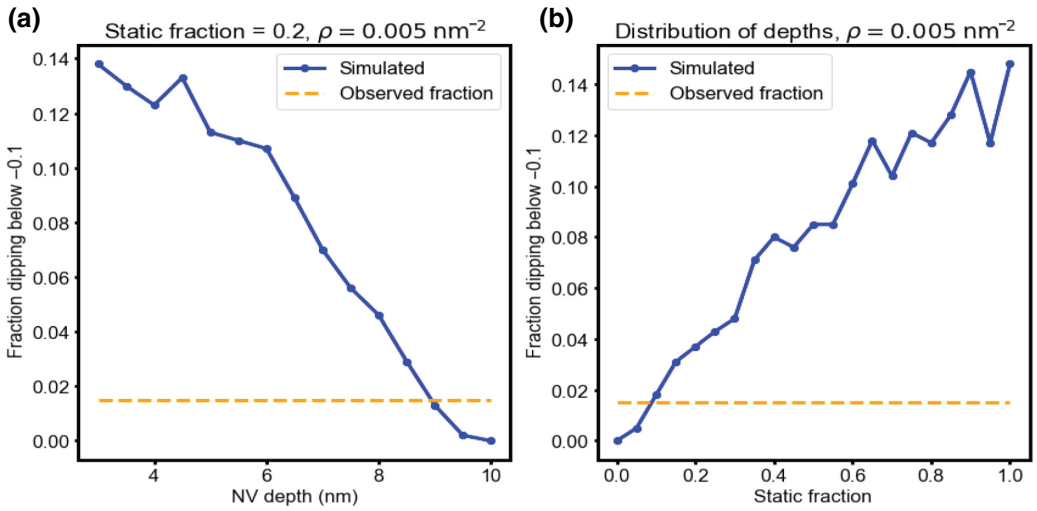


FIG. 22. Stationary surface spin fraction. (a) Simulation of the fraction of NV centers that exhibit strong coupling in DEER as a function of depth, if 20% of spins on the surface are placed randomly relative to the NV center and then do not hop. (b) Simulation of the fraction of NV centers that exhibit strong coupling in their DEER signal as a function of the fraction of spins that are stationary on the surface, given a density of $\sigma = 0.005$ nm^{-2}, and depths drawn from the distribution of depths measured across all samples. Orange dashed line corresponds to the observed fraction of three strongly coupled DEER signals out of 191 NV centers.

expected to take in either the static spin or configurational averaging picture. To assess the utility of the configurational averaging model, we compare the fit of the Monte Carlo simulated spin hopping model to the single exponential model fit for the four NV centers in Fig. 20 (three of which are also depicted in Fig. 3 of the main text) using several standard Bayesian model comparisons [59,60]. For each of these NV centers, we report the widely applicable information criteria (WAIC) and associated error, σ_{WAIC}, the leave-one-out (LOO) cross-validation statistic and associated error, σ_{LOO}, and the total model log-likelihood and associated error, $\sigma_{\mathrm{Log\text{-}likelihood}}$, for the two different models. The results are reported in Table II, from which it is apparent that the configurational averaging model is preferred over the single exponential predicted from a static bath in all cases, and that this difference is significant for two NV centers. The reason for the low significance of NV1 is likely the lower signal to noise of this measurement, and for NV4, the two models are expected to perform similarly due to the depth of this NV center.

For the single exponential fit, the stretching factor n is constrained to be $n \geq 1$ following the discussion of noise sources in Appendix F 1. For the configurational averaging Monte Carlo model, both the depth and the surface spin density are free parameters. In order to remove any ambiguity regarding data cutoffs, the coherence and the DEER signal are fit simultaneously with separate Gaussian likelihoods. This is possible for these NV centers because the host diamond is isotopically purified, and the absence of ^{13}C oscillations greatly simplifies the fitting. The coherence is therefore assumed to take the form of a stretched single exponential.

APPENDIX I: FRACTION OF SPINS THAT ARE HOPPING

Across the nine samples listed in Table I of Appendix A, and an additional five samples with similar implantation parameters and surface preparations, we measure DEER signals on 191 unique NV centers and observe only three spins that exhibit obvious coherent oscillations, as defined by a DEER contrast (DEER echo divided by coherence) that dip below -0.1. This ratio allows us to place a bound on the number of spins that could still be stationary while the remaining spins hop. If we assume typical extracted densities of $\sigma = 0.005$ nm^{-2} then, given that the typical NV center depths in our samples are on average less than 10 nm, we expect that less than 20% of surface spins are stationary (see Fig. 22).

[1] Jeronimo R. Maze, Paul L. Stanwix, James S. Hodges, Seungpyo Hong, Jacob M. Taylor, Paola Cappellaro, Liang Jiang, M. V. Gurudev Dutt, Emre Togan, and A. S. Zibrov, *et al.*, Nanoscale magnetic sensing with an individual electronic spin in diamond, Nature **455**, 644 (2008).

[2] Jianming Cai, Alex Retzker, Fedor Jelezko, and Martin B. Plenio, A large-scale quantum simulator on a diamond surface at room temperature, Nat. Phys. **9**, 168 (2013).

[3] Lukas Schlipf, Thomas Oeckinghaus, Kebiao Xu, Durga Bhaktavatsala Rao Dasari, Andrea Zappe, Felipe Favaro De Oliveira, Bastian Kern, Mykhailo Azarkh, Malte Drescher, and Markus Ternes, *et al.*, A molecular quantum spin network controlled by a single qubit, Sci. Adv. **3**, e1701116 (2017).

[4] Bernhard Grotz, Johannes Beck, Philipp Neumann, Boris Naydenov, Rolf Reuter, Friedemann Reinhard, Fedor Jelezko, Jörg Wrachtrup, David Schweinfurth, Biprajit Sarkar, and Philip Hemmer, Sensing external spins with nitrogen-vacancy diamond, New J. Phys. **13**, 055004 (2011).

[5] M. S. Grinolds, M. Warner, K. De Greve, Y. Dovzhenko, L. Thiel, R. L. Walsworth, S. Hong, P. Maletinsky, and A. Yacoby, Subnanometre resolution in three-dimensional magnetic resonance imaging of individual dark spins, Nat. Nanotechnol. **9**, 279 (2014).

[6] I. Lovchinsky, A. O. Sushkov, E. Urbach, N. P. de Leon, S. Choi, K. De Greve, R. Evans, R. Gertner, E. Bersin, C. Müller, L. McGuinness, F. Jelezko, R. L. Walsworth, H. Park, and M. D. Lukin, Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic, Science (New York, N.Y.) **351**, 836 (2016).

[7] A. O. Sushkov, I. Lovchinsky, N. Chisholm, R. L. Walsworth, H. Park, and M. D. Lukin, Magnetic Resonance Detection of Individual Proton Spins Using Quantum Reporters, Phys. Rev. Lett. **113**, 197601 (2014).

[8] B. A. Myers, A. Das, M. C. Dartiailh, K. Ohno, D. D. Awschalom, and A. C. Bleszynski Jayich, Probing Surface Noise with Depth-Calibrated Spins in Diamond, Phys. Rev. Lett. **113**, 027602 (2014).

[9] Y. Romach, C. Müller, T. Unden, L. J. Rogers, T. Isoda, K. M. Itoh, M. Markham, A. Stacey, J. Meijer, S. Pezzagna, B. Naydenov, L. P. McGuinness, N. Bar-Gill, and F. Jelezko, Spectroscopy of Surface-Induced Noise Using Shallow Spins in Diamond, Phys. Rev. Lett. **114**, 017601 (2015).

[10] Sorawis Sangtawesin, Bo L. Dwyer, Srikanth Srinivasan, James J. Allred, Lila V. H. Rodgers, Kristiaan De Greve, Alastair Stacey, Nikolai Dontschuk, Kane M. O'Donnell, and Di Hu, *et al.*, Origins of Diamond Surface Noise Probed by Correlating Single-Spin Measurements with Surface Spectroscopy, Phys. Rev. X **9**, 031052 (2019).

[11] H. J. Mamin, M. H. Sherwood, and D. Rugar, Detecting external electron spins using nitrogen-vacancy centers, Phys. Rev. B **86**, 195422 (2012).

[12] Dolev Bluvstein, Zhiran Zhang, Claire A. McLellan, Nicolas R. Williams, and Ania C. Bleszynski Jayich, Extending the Quantum Coherence of a near-Surface Qubit by Coherently Driving the Paramagnetic Surface Environment, Phys. Rev. Lett. **123**, 146804 (2019).

[13] Linh M. Pham, Stephen J. DeVience, Francesco Casola, Igor Lovchinsky, Alexander O. Sushkov, Eric Bersin, Junghyun Lee, Elana Urbach, Paola Cappellaro, Hongkun Park, Amir Yacoby, Mikhail Lukin, and Ronald L. Walsworth, NMR technique for determining the depth of shallow nitrogen-vacancy centers in diamond, Phys. Rev. B **93**, 045425 (2016).

[14] A. D. Milov, K. M. Salikhov, and M. D. Shirov, Application of the double resonance method to electron spin echo in a study of the spatial distribution of paramagnetic centers in solids, Sov. Phys. Solid State **23**, 565 (1981).

[15] L. Childress, M. V. Gurudev Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, Coherent dynamics of coupled electron and nuclear spin qubits in diamond, Science **314**, 281 (2006).

[16] L._G. Rowan, E. L. Hahn, and W. B. Mims, Electron-spin-echo envelope modulation, Phys. Rev. **137**, A61 (1965).

[17] M. Kim, H. J. Mamin, M. H. Sherwood, K. Ohno, D. D. Awschalom, and D. Rugar, Decoherence of Near-Surface Nitrogen-Vacancy Centers Due to Electric Field Noise, Phys. Rev. Lett. **115**, 087602 (2015).

[18] Rogerio de Sousa, in *Electron Spin Resonance and Related Phenomena in Low-Dimensional Structures*, edited by Marco Fanciulli (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009), p. 183.

[19] J. R. Klauder and P. W. Anderson, Spectral diffusion decay in spin resonance experiments, Phys. Rev. **125**, 912 (1962).

[20] V. V. Dobrovitski, A. E. Feiguin, D. D. Awschalom, and R. Hanson, Decoherence dynamics of a single spin versus spin ensemble, Phys. Rev. B **77**, 245212 (2008).

[21] P. L. Stanwix, L. M. Pham, J. R. Maze, D. Le Sage, T. K. Yeung, P. Cappellaro, P. R. Hemmer, A. Yacoby, M. D. Lukin, and R. L. Walsworth, Coherence of nitrogen-vacancy electronic spin ensembles in diamond, Phys. Rev. B **82**, 201201 (2010).

[22] P. Siyushev, H. Pinto, M. Vörös, A. Gali, F. Jelezko, and J. Wrachtrup, Optically Controlled Switching of the Charge State of a Single Nitrogen-Vacancy Center in Diamond at Cryogenic Temperatures, Phys. Rev. Lett. **110**, 167402 (2013).

[23] Janik Wolters, Nikola Sadzak, Andreas W. Schell, Tim Schröder, and Oliver Benson, Measurement of the Ultrafast Spectral Diffusion of the Optical Transition of Nitrogen Vacancy Centers in Nano-Size Diamond Using Correlation Interferometry, Phys. Rev. Lett. **110**, 027401 (2013).

[24] L. C. Bassett, F. J. Heremans, C. G. Yale, B. B. Buckley, and D. D. Awschalom, Electrical Tuning of Single Nitrogen-Vacancy Center Optical Transitions Enhanced by Photoinduced Fields, Phys. Rev. Lett. **107**, 266403 (2011).

[25] Yiwen Chu, Nathalie Pulmones de Leon, Brendan J. Shields, Birgit Hausmann, R. Evans, Emre Togan, Michael John Burek, M. Markham, Alastair Stacey, and Alexander S. Zibrov, *et al.*, Coherent optical transitions in implanted nitrogen vacancy centers, Nano Lett. **14**, 1982 (2014).

[26] Serge Lacelle and Luc Tremblay, What is a typical dipolar coupling constant in a solid?, J. Chem. Phys. **98**, 3642 (1993).

[27] Edward B. Fel'dman and Serge Lacelle, Configurational averaging of dipolar interactions in magnetically diluted spin networks, J. Chem. Phys. **104**, 2000 (1996).

[28] Walter Hahn and V. V. Dobrovitski, Long-lived coherence in driven spin systems: from two-to infinite spatial dimensions, (2019), arXiv preprint ArXiv:1911.06272.

[29] Felipe Fávaro de Oliveira, Denis Antonov, Ya Wang, Philipp Neumann, Seyed Ali Momenzadeh, Timo Häußermann, Alberto Pasquarelli, Andrej Denisenko, and Jörg Wrachtrup, Tailoring spin defects in diamond by lattice charging, Nat. Commun. **8**, 15409 (2017).

[30] Yuki Doi, Takahiro Fukui, Hiromitsu Kato, Toshiharu Makino, Satoshi Yamasaki, Toshiyuki Tashima, Hiroki Morishita, Shinji Miwa, Fedor Jelezko, and Yoshishige Suzuki, *et al.*, Pure negatively charged state of the NV center in *n*-type diamond, Phys. Rev. B **93**, 081203 (2016).

[31] Brendon C. Rose, Ding Huang, Zi-Huai Zhang, Paul Stevenson, Alexei M. Tyryshkin, Sorawis Sangtawesin,

Srikanth Srinivasan, Lorne Loudin, Matthew L. Markham, and Andrew M. Edmonds, *et al.*, Observation of an environmentally insensitive solid-state spin defect in diamond, Science **361**, 60 (2018).

[32] E. D. Herbschleb, H. Kato, Y. Maruyama, T. Danjo, T. Makino, S. Yamasaki, I. Ohki, K. Hayashi, H. Morishita, and M. Fujiwara, *et al.*, Ultra-long coherence times amongst room-temperature solid-state spins, Nat. Commun. **10**, 1 (2019).

[33] N. J. Glaser, G. Braunbeck, O. Bienek, I. D. Sharp, and F. Reinhard, Can surface-transfer doping and UV irradiation during annealing improve shallow implanted nitrogen-vacancy centers in diamond?, Appl. Phys. Lett. **117**, 054003 (2020).

[34] Moloud Kaviani, Peter Deák, Bálint Aradi, Thomas Frauenheim, Jyh-Pin Chou, and Adam Gali, Proper surface termination for luminescent near-surface NV centers in diamond, Nano Lett. **14**, 4772 (2014).

[35] M. S. Grinolds, P. Maletinsky, Sungkun Hong, M. D. Lukin, R. L. Walsworth, and A. Yacoby, Quantum control of proximal spins using nanoscale magnetic resonance imaging, Nat. Phys. **7**, 687 (2011).

[36] Haitao Liu, Sunmin Ryu, Zheyuan Chen, Michael L. Steigerwald, Colin Nuckolls, and Louis E. Brus, Photochemical reactivity of graphene, J. Am. Chem. Soc. **131**, 17099 (2009).

[37] Oleg V. Yazyev and Lothar Helm, Defect-induced magnetism in graphene, Phys. Rev. B **75**, 125408 (2007).

[38] Lucio Robledo, Lilian Childress, Hannes Bernien, Bas Hensen, Paul F. A. Alkemade, and Ronald Hanson, High-fidelity projective read-out of a solid-state spin quantum register, Nature **477**, 574 (2011).

[39] Dominik M. Irber, Francesco Poggiali, Fei Kong, Michael Kieschnick, Tobias Lühmann, Damian Kwiatkowski, Jan Meijer, Jiangfeng Du, Fazhan Shi, and Friedemann Reinhard, Robust all-optical single-shot readout of NV centers in diamond, (2020), ArXiv:2006.02938.

[40] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui, Single spin detection by magnetic resonance force microscopy, Nature **430**, 329 (2004).

[41] Patrick Maletinsky, Sungkun Hong, Michael Sean Grinolds, Birgit Hausmann, Mikhail D. Lukin, Ronald L. Walsworth, Marko Loncar, and Amir Yacoby, A robust scanning diamond sensor for nanoscale imaging with single nitrogen-vacancy centres, Nat. Nanotechnol. **7**, 320 (2012).

[42] Matthew Pelliccione, Alec Jenkins, Preeti Ovartchaiyapong, Christopher Reetz, Eve Emmanouilidou, Ni Ni, and Ania C. Bleszynski Jayich, Scanned probe imaging of nanoscale magnetism at cryogenic temperatures with a single-spin quantum sensor, Nat. Nanotechnol. **11**, 700 (2016).

[43] Emily J. Davis, Bingtian Ye, Francisco Machado, Simon A. Meynell, Thomas Mittiga, William Schenken, Maxime Joos, Bryce Kobrin, Yuanqi Lyu, and Dolev Bluvstein, *et al.*, Probing many-body noise in a strongly interacting two-dimensional dipolar spin system, (2021), arXiv preprint ArXiv:2103.12742.

[44] J. Stöhr, Springer series in surface science, NEXAFS Spectrosc. (Springer, Berlin, Heidelberg, 1992), Vol. 25, p. 114.

[45] K. Bobrov, G. Comtet, G. Dujardin, and L. Hellner, Electronic Structure of Partially Hydrogenated Si (100)-(2×1) Surfaces Prepared by Thermal and Nonthermal Desorption, Phys. Rev. Lett. **86**, 2633 (2001).

[46] Zhi-Hui Wang, G. De Lange, D. Ristè, R. Hanson, and V. V. Dobrovitski, Comparison of dynamical decoupling protocols for a nitrogen-vacancy center in diamond, Phys. Rev. B **85**, 155204 (2012).

[47] D. J. Chadi, Stabilities of single-layer and Bilayer Steps on Si (001) Surfaces, Phys. Rev. Lett. **59**, 1691 (1987).

[48] H. X. Yang, L. F. Xu, Z. Fang, C. Z. Gu, and S. B. Zhang, Bond-Counting Rule for Carbon and Its Application to the Roughness of Diamond (001), Phys. Rev. Lett. **100**, 026101 (2008).

[49] Georg Kresse and Jürgen Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[50] Peter E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[51] John P. Perdew, Kieron Burke, and Yue Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system, Phys. Rev. B **54**, 16533 (1996).

[52] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof, Hybrid functionals based on a screened Coulomb potential, J. Chem. Phys. **118**, 8207 (2003).

[53] Adam Gali, Erik Janzén, Péter Deák, Georg Kresse, and Efthimios Kaxiras, Theory of Spin-Conserving Excitation of the N-V^- Center in Diamond, Phys. Rev. Lett. **103**, 186404 (2009).

[54] Charles P. Slichter, *Principles of Magnetic Resonance* (Springer-Verlag, Berlin, Heidelberg, New York, 1990), 655.

[55] Anatole Abragam, *The Principles of Nuclear Magnetism* (Oxford university press, Oxford, England, 1961).

[56] Abdelghani Laraoui, Florian Dolde, Christian Burk, Friedemann Reinhard, Jörg Wrachtrup, and Carlos A. Meriles, High-resolution correlation spectroscopy of 13 C spins near a nitrogen-vacancy centre in diamond, Nat. Commun. **4**, 1 (2013).

[57] B. A. Myers, A. Ariyaratne, and A. C. Bleszynski Jayich, Double-Quantum Spin-Relaxation Limits to Coherence of Near-Surface Nitrogen-Vacancy Centers, Phys. Rev. Lett. **118**, 197201 (2017).

[58] L. T. Hall, P. Kehayias, D. A. Simpson, A. Jarmola, A. Stacey, D. Budker, and L. C. L. Hollenberg, Detection of nanoscale electron spin resonance spectra demonstrated using nitrogen-vacancy centre probes in diamond, Nat. Commun. **7**, 10211 (2016).

[59] Aki Vehtari, Andrew Gelman, and Jonah Gabry, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Stat. Comput. **27**, 1413 (2017).

[60] Phil Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support* (Cambridge University Press, Cambridge, England, 2005).