Supplemental Materials: Ultralong dephasing times in solid-state spin ensembles via quantum control

Erik Bauch,[1,*] Connor A. Hart,[1,*] Jennifer M. Schloss,[1,2,3] Matthew J. Turner,[1,3] John F. Barry,[1,4] Pauli Kehayias,[1,5] Swati Singh,[6] and Ronald L. Walsworth[1,3,5,†]

[1] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[2] *Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3] *Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA*
[4] *Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420, USA*
[5] *Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA*
[6] *Williams College, Department of Physics, 33 Lab Campus Drive, Williamstown, Massachusetts 01267*

CONTENTS

I. Experimental methods 2

II. Sample information (all samples) 3

III. Survey of dephasing times 3

IV. Strain contribution to T_2^* 3

V. Spin bath contribution to T_2^* 3

VI. Magnetic field gradient contribution to T_2^* 5

VII. NV Hamiltonian in single and double quantum bases 6
Case 1: Zero strain, zero off-axis magnetic field 6
Case 2: Non-zero strain, zero off-axis magnetic field 7
Case 3: Non-zero off-axis magnetic field 7

VIII. Spin bath driving model 7

IX. ^{14}N and ^{15}N double electron-electron resonance spectra 8
^{14}N spectrum 9
^{15}N spectrum 10

X. Continuous versus pulsed spin bath driving 10

XI. NV and nitrogen spin resonance linewidth measurements 11

XII. DC magnetometry with DQ and spin-bath drive 13

XIII. Dephasing channels per sample 16

References 18

* these authors contributed equally to this work

† rwalsworth@cfa.harvard.edu

I. EXPERIMENTAL METHODS

A custom-built, wide-field microscope collected the spin-dependent fluorescence from an NV ensemble onto an avalanche photodiode. Optical initialization and readout of the NV ensemble was accomplished via 532 nm continuous-wave (CW) laser light focused through the same objective used for fluorescence collection (Fig. 1a). The detection volume was given by the 532 nm beam excitation at the surface (diameter $\approx 20\,\mu$m) and sample thickness (100 µm for Samples A and B, 40 µm for Sample C). A static magnetic bias field was applied to split the $|-1\rangle$ and $|+1\rangle$ degeneracy in the NV ground state using two permanent samarium cobalt ring magnets in a Helmholtz-type configuration, with the generated field aligned along one [111] crystallographic axis of the diamond ($\equiv \hat{z}$). The magnet geometry was optimized using the Radia software package [S1] to minimize field gradients over the detection volume (see Suppl. VI). A planar waveguide fabricated onto a glass substrate delivered $2-3.5$ GHz microwave radiation for coherent control of the NV ensemble spin states. To manipulate the nitrogen spin resonances (see Fig. 1d), a 1 mm-diameter copper loop was positioned above the diamond sample to apply $100-600$ MHz radiofrequency (RF) signals, synthesized from up to eight individual signal generators. Pulsed measurements on the NV and nitrogen spins were performed using a computer-controlled pulse generator and microwave switches.

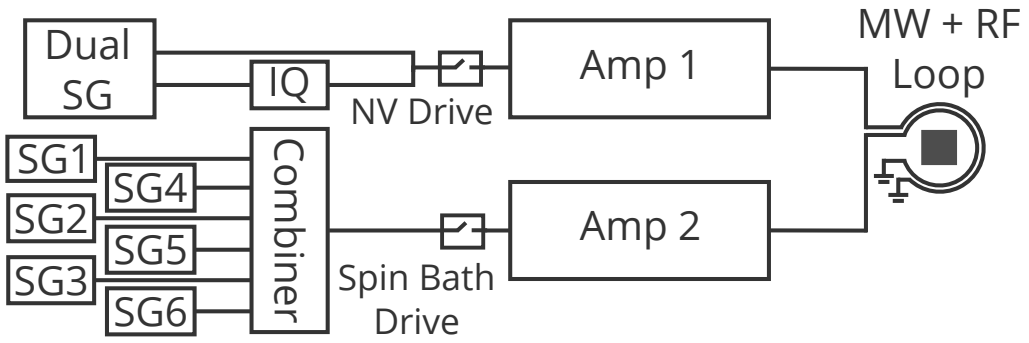


Figure S1. Microwave generation and delivery schematic. For NV spin state control: Single and two-tone signals are generated using a dual channel Windfreak Technology Synth HD signal generator. One channel includes a Marki IQ-1545 mixer to manipulate the relative phase between both channels. A single Minicircuits ZASWA-2-50DR+ switch is used to generate the NV control pulses before amplification with a Minicircuits ZHL-16W-43 amplifer. The NV control fields are delivered to the diamond sample using a fabricated microwave waveguide (diameter 500 µm). For spin bath control: Up to eight single channel Windfreak Technology Synth NV signal generators are combined before passing through a switch and a Minicircuits ZHL-100W-52 100 W amplifier. The amplified field is delivered via a grounded cooper loop (1 mm diameter).

The NV ESR measurement contrast (Fig. 2, 3, and 4d) is determined by comparing the fluorescence from the NV ensemble in the $|0\rangle$ state (maximal fluorescence) relative to the $|+1\rangle$ or $|-1\rangle$ state (minimal fluorescence) [S2] and is defined as visibility $C = \frac{\text{max}-\text{min}}{\text{max}+\text{min}}$. The DEER (Fig. 1d) and DC magnetometry contrast (Fig. 4d) are calculated in the same fashion, but are reduced by $\approx 1/e$ since the best phase sensitivity in those measurements is obtained at $\tau \approx T_2$ and $\tau \approx T_2^*$, respectively (see Suppl. XI and XII). For noise rejection, most pulse sequences in this work use a back-to-back double measurement scheme [S3], where the accumulated NV spin ensemble phase signal is first projected onto the $|0\rangle$ state and then onto the $|+1\rangle$ (or $|-1\rangle$) state. The contrast for a single measurement is then defined as the visibility of both sequences.

II. SAMPLE INFORMATION (ALL SAMPLES)

Information for all samples used in this study is summarized in Table S1.

Table S1. Detailed information for Samples A - E. Values with $\sim$ symbol are order-of-magnitude estimates. For all samples, [NV] $\ll$ [N] and NV contributions to T_2^* can be neglected ($1\,\text{ppm} = 1.76 \times 10^{17}\,\text{cm}^{-3}$).

Sample	[N] (ppm)	^{13}C (%)	[NV] (cm^{-3})	T_2^{meas} (μs)	$T_{2,\text{SQ}}^{*,\text{meas}}$ (μs)	$T_{2,\text{DQ}}^{*,\text{meas}}$ (μs)	$T_{2,\text{NV-N}}^{*,\text{est}}$ (μs)	$T_{2,\text{NV-13C}}^{*,\text{est}}$ (μs)	$T_{2,\text{NV-(13C+N)}}^{*,\text{est}}$ (μs)	dM_z^{meas}/dL (MHz/μm)
A	$\lesssim 0.05$	0.01	$\sim 3 \times 10^{12}$	$\gtrsim 630$	$5-12$	34	350	100	78	n/a
B	0.75	0.01	$\sim 10^{14}$	$250-300$	$1-10$	6.9	23	100	19	0.0028
C	10	0.05	$\sim 6 \times 10^{15}$	15-18	$0.3-1.2$	0.6	2	20	2	n/a
D	3	< 0.01	$\sim 5 \times 10^{15}$	53	0.4	1.3	6	100	> 6	n/a
E	48	1.1	$\sim 1 \times 10^{17}$	1.8	0.07	0.12	0.3	1	0.2	n/a

III. SURVEY OF DEPHASING TIMES

In Fig. S2 we show a survey of inhomogeneous dephasing times for electronic solid-state spin ensembles.

IV. STRAIN CONTRIBUTION TO T_2^*

The on-axis strain component M_z in Sample B was mapped across a $1 \times 1\,\text{mm}$ area using a separate wide-field imager of NV spin-state-dependent fluorescence. A bias field $B_0 \sim 1.5\,\text{mT}$ was applied to split the spin resonances from the four NV orientations. Measurements were performed following the vector magnetic microscopy (VMM) technique [S13]. Eqn. 3 in the main text was used to analyze the measured NV resonance frequencies from each camera pixel (ignoring M_x and M_y terms as small perturbations, see Suppl. VII). This procedure yielded the average B_x, B_y, and B_z magnetic field components, as well as the M_z on-axis strain components for all four NV orientations in each camera pixel, corresponding to $2.42\,\mu\text{m} \times 2.42\,\mu\text{m}$ transverse resolution on the diamond sample. Figure 1c of the main text shows the resulting map of the on-axis strain inhomogeneity M_z in Sample B for the NV orientation interrogated in this work. This map indicates an approximate strain gradient of 2.8 kHz/μm across the field of view. The estimated strain gradient was used for all samples, while recognizing the likely variation between samples and within different regions of a sample. Across a 20-μm diameter spot, the measured strain inhomogeneity corresponds to a T_2^* limit of $\approx 6\,\mu\text{s}$, which compares well with the measured variation in $T_{2,\text{SQ}}^*$ for Samples A and B (see Table 1). Note that the contributions to M_z can be microscopic (e.g., due to nearby point defects) or macroscopic (e.g., due to crystal defects with size $> 10\,\mu\text{m}$). In addition, the VMM technique integrates over macroscopic gradients within the depth of field of the VMM microscope. For the present experiments the resolution along the z-axis (i.e., perpendicular to the diamond surface) is given approximately by the thickness of the NV-diamond layer. Consequently, the strain gradient estimate shown in Fig. 1c is a measure of M_z gradients in-plane within the NV layer, and strain gradients across the NV layer thickness are not resolvable in this measurement.

V. SPIN BATH CONTRIBUTION TO T_2^*

The NV spin ensemble T_2^* as a function of nitrogen concentration is estimated from the average dipolar coupling between electronic nitrogen spins, which is given by $\gamma_{e-e} = a \times \frac{\mu_0}{4\pi} g^2 \mu_B^2 / \hbar \frac{1}{\langle r \rangle^3} \approx 2\pi \times 9.1 \cdot [\text{N}]\,\text{kHz/ppm}$, where μ_0 is the vacuum permeability, g is the electron g-factor, μ_B is the Bohr magneton, $\hbar$ is the reduced Planck constant, $\langle r \rangle = 0.55[\text{N}]^{-1/3}$ is the average spacing between electronic nitrogen spins as a function of density [N] (in parts-per-million) within diamond [S14], and a is a factor of order unity collecting additional parameters from the dipolar estimate such as the angular dependence and spin resonance lineshape of the ensemble [S15]. A sample with $[\text{N}] = 1\,\text{ppm}$ has an estimated $T_{2,\text{NV-N}}^* \approx 1/(2\pi \times 9.1\,\text{kHz}) = 17.5\,\mu\text{s}$ using this dipolar estimate. Similarly, Table S1 gives the estimates $T_{2,\text{NV-N}}^*$ for Samples A, B, and C.

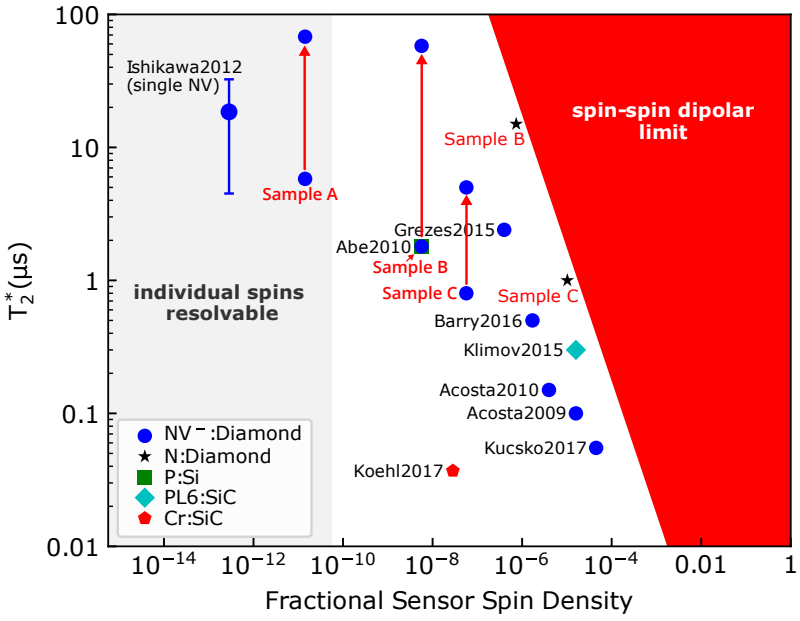


Figure S2. Inhomogeneous spin dephasing times. Experimental results from this work are compared to that of related spin defect systems (see legend). Inhomogeneous dephasing due to paramagnetic bath spins (e.g., nitrogen and ^{13}C nuclear spins in diamond), strain fields and other effects limit $T_{2\,\mathrm{ens}}^*$ at lower sensor-spin densities $\ll 1$. At higher sensor-spin densities approaching unity, spin-spin interaction places an upper bound on the ensemble dephasing time (red shaded area). This limit to T_2^* is estimated using γ_{e-e} (see Suppl. V) and a fractional sensor-spin density of 1 corresponds to $\sim 10^{23}\,\mathrm{cm}^{-3}$. Red arrows indicate improvement from the bare T_2^* as measured in the NV SQ basis and increase when DQ sensing and spin bath drive (where advantageous) are employed to suppress inhomogeneities. The maximal obtained T_2^* values for Sample A, B and C are multiplied by a factor of two to account for the twice higher gyromagnetic ratio in the NV DQ basis. The region in which individual, single spins are resolvable with confocal microscopy ($\sim 200\,\mathrm{nm}$ average spin separation) is shown in gray. T_2^* values determined in Sec. XI for the nitrogen spins (P1 centers) are shown for reference as well (black stars). Measurements by Ishikawa et al. were performed on single NV centers and error bars indicate the spread in measured T_2^* values. The following references were used: Abe2010 [S4], Acosta2009 [S5], Acosta2010 [S6], Barry2016 [S7], Grezes2015 [S8], Ishikawa2012 [S9], Klimov2015 [S10], Koehl2017 [S11], and Kucsko2017 [S12].

Uncertainties in nitrogen concentration [N] used in Fig. 4c are estimated by considering: the values reported by the manufacturer (Element Six Inc.); fluorescence measurements in a confocal microscope (Sample A); and Hahn echo T_2 measurements using the calibration value $T_2(\mathrm{N}) \simeq 165\,\mu\mathrm{s}\cdot\mathrm{ppm}$ reported in Ref. [S16] (Samples B and C). For example, for Sample B, Element Six reports $[\mathrm{N}] = 1\,\mathrm{ppm}$, whereas the measured $T_2 = 300\,\mu\mathrm{s}$ suggests $[\mathrm{N}] = 0.5\,\mathrm{ppm}$. The average value is thus used: $[\mathrm{N}] = 0.75 \pm 0.25\,\mathrm{ppm}$.

In the dilute ^{13}C limit ($n_{13\mathrm{C}} \lesssim 1.1\,\%$, where $n_{13\mathrm{C}}$ is the ^{13}C spin concentration in percent), the NV-^{13}C contact interaction can be neglected and thus the NV ensemble ESR linewidth is expected to be linearly-dependent on the ^{13}C concentration [S4, S15], i.e., $1/T_{2,\mathrm{NV\text{-}13C}}^* = A_{\mathrm{NV\text{-}^{13}C}} \cdot n_{13\mathrm{C}}$. An NV spin ensemble T_2^* measurement on a natural abundance sample with $n_{13\mathrm{C}} = 1.07\,\%$ therefore provides a reasonable lower-bound estimate for $A_{\mathrm{NV\text{-}13C}}$ from which the ^{13}C contribution in our diamond samples can be calculated. Fig. S3 shows a DQ Ramsey measurement of a natural ^{13}C abundance sample. Via a fit to the Ramsey data in the time domain, we extract $T_{2,\mathrm{DQ}}^* = 445(30)\,\mathrm{ns}$ and $p = 1.0(1)$. After correcting for the small contribution of 0.4 ppm nitrogen spins in the sample using the calibration found in Fig. 4c of the main text, we calculate $A_{\mathrm{NV\text{-}13C}} \approx 2\pi \times 160\,\mathrm{kHz}/\%$ ($1/A_{^{13}\mathrm{C}} \approx 1\,\mu\mathrm{s}\cdot\%$) from which we determine the NV-^{13}C limits given in Table 1 and the main text of the paper.

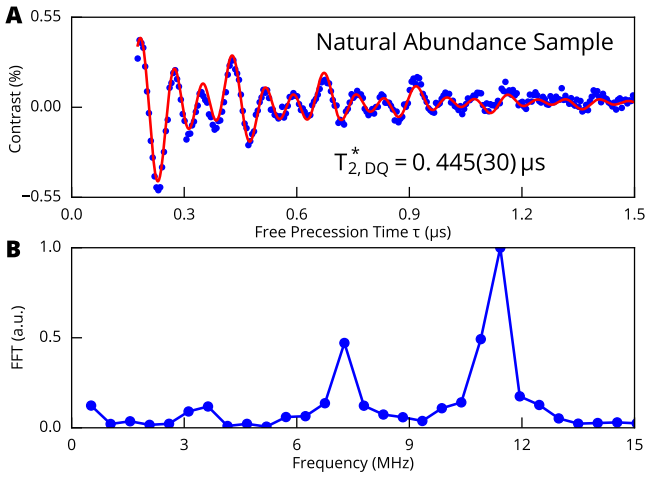


Figure S3. NV Ramsey measurement for natural isotope abundance diamond sample. (a) DQ Ramsey measurement on a natural abundance sample ([N] $\simeq$ 0.4 ppm, [^{13}C] = 1.07 %) yields $T^*_{2,\mathrm{DQ}}$ = 0.445(30) µs. (b) Fourier transform of Ramsey signal showing the enhanced precession in the DQ basis. A frequency detuning from the center hyperfine state of 3.65 MHz was chosen in this measurement; by sensing in the DQ basis, the detuning from each hyperfine state has acquired a factor of two.

VI. MAGNETIC FIELD GRADIENT CONTRIBUTION TO T_2^*

The NV-diamond epifluorescence microscope employs a custom-built samarium-cobalt (SmCo) magnet geometry designed to apply a homogeneous external field B_0 parallel to NVs oriented along the [111] diamond crystallographic axis. The field strength can be varied from 2 to 20 mT (Fig. S4a). SmCo was chosen for its low reversible temperature coefficient (-0.03 %/K). Calculations performed using the Radia software package [S1] enabled the optimization of the geometry to minimize B_0 gradients across the NV fluorescence collection volume. This collection volume is approximately cylindrical, with a measured diameter of $\approx$ 20 µm and a length determined by the NV layer thickness along the z-axis (40 − 100 µm, depending on the diamond sample; see descriptions in the main text).

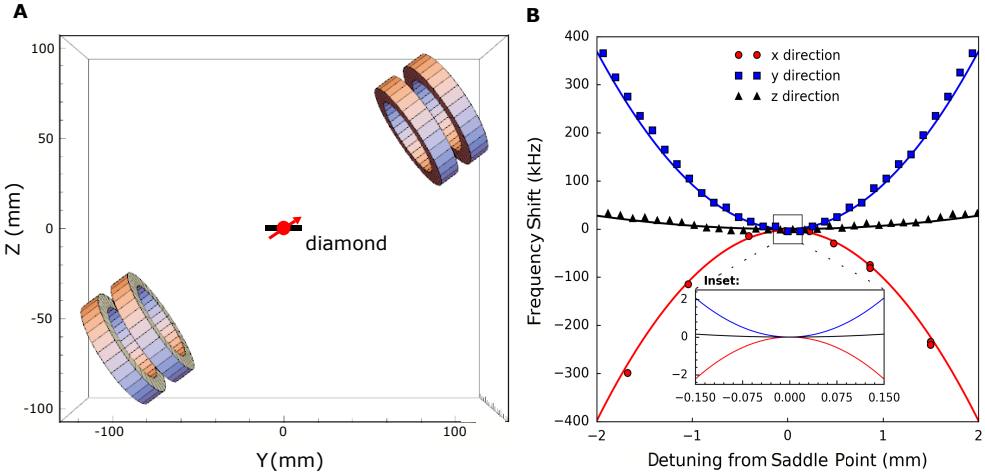


Figure S4. Design of homogeneous magnetic bias field. (a) Magnet geometry used to apply an external B_0 field along one NV orientation within the diamond crystal (typically [111]) as modeled using Radia [S1]. Red arrow depicts the NV orientation class interrogated in these experiments; black rectangle represents diamond sample approximately to scale. (b) Magnets are translated along three axes to measure the B_0 field strength (shift in ESR transition frequency) as a function of detuning from the origin $(x, y, z = 0)$ where the origin is defined as the center of the collection volume. Solid lines depict Radia simulation results while plotted points correspond to measured values. Inset: Zoomed-in view for length scale relevant for NV fluorescence collection volumes used in this work.

To calculate the expected B_0 field strength along the target NV orientation, the dimensions and properties of the magnets were used as Radia input, as well as an estimated 3° misalignment angle of the magnetic field with the NV axis. We find good agreement between the calculated field strength and values extracted from NV ESR measurements in Sample B, over a few millimeter lengthscale. The simulation results and measured values are plotted together in Fig. S4b. The z-direction gradient is reduced compared to the gradient in the xy-plane due to a high degree of symmetry along the z-axis for the magnet geometry.

Using data and simulation, we calculate that the B_0 gradient at 8.5 mT induces an NV ensemble ESR linewidth broadening of less than 0.1 kHz across the collection volume of Sample B. This corresponds to a T_2^*-limit on the order of 1 ms. However, due to interaction of the bias magnetic field with nearby materials and the displacement of the collection volume from the magnetic field saddle point, the experimentally realized gradient for Sample B was found to contribute an NV ESR linewidth broadening ≈ 1 kHz (implying a T_2^*-limit $\approx 320\,\mu$s), which constitutes a small but non-negligible contribution to the T_2^* values measured in this work. Ramsey measurements for Sample A were taken at a four times smaller bias field; we estimate therefore $\approx 4\times$ better magnetic field homogeneity. For Sample C, with a layer thickness of 40 µm, the contribution of the magnetic field gradient at 10 mT to T_2^* was similar to that of Sample B.

VII. NV HAMILTONIAN IN SINGLE AND DOUBLE QUANTUM BASES

In this section we discuss the influence of strain and magnetic fields in the single quantum (SQ) and double quantum (DQ) bases by considering several limiting cases. We first discuss how common-mode noise sources, i.e., sources that shift the NV $|-1\rangle$ and $|+1\rangle$ energy levels in-phase and with equal magnitude, are suppressed in the DQ basis. We then discuss how off-NV-axis strain fields are suppressed even by moderate bias magnetic fields. Lastly, we discuss the effect of off-axis magnetic fields on the NV spin-state energy levels and T_2^*. We begin with the negatively-charged NV ground electronic state electronic spin ($S = 1$) Hamiltonian, which is given by [S2] (neglecting hyperfine and quadrupolar effects):

$$\mathbf{H}/h = D\mathbf{S}_z^2 + \frac{\gamma_{\mathrm{NV}}}{2\pi}(B_x\mathbf{S}_x + B_y\mathbf{S}_y + B_z\mathbf{S}_z) + M_z\mathbf{S}_z^2 + M_x(\mathbf{S}_y^2 - \mathbf{S}_x^2) + M_y(\mathbf{S}_x\mathbf{S}_y + \mathbf{S}_y\mathbf{S}_x), \tag{S1}$$

where $D \approx 2.87$ GHz is the NV zero-field splitting due to spin-spin interactions, $\{B_x, B_y, B_z\}$ are the magnetic field components, $\{M_x, M_y, M_z\}$ collect strain and electric field components, $\{\mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z\}$ are the dimensionless spin-1 operators, and $\frac{g\mu_B}{h} = \frac{\gamma_{\mathrm{NV}}}{2\pi} \approx 28.025$ GHz/T is the NV gyromagnetic ratio. Using $M_\perp \equiv -(M_x + iM_y)$, $B_\perp \equiv \frac{1}{\sqrt{2}}(B_x + iB_y)$, and the standard definitions for the spin operators $\{\mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z\}$, Eqn. S1 reads in matrix form:

$$\mathbf{H}/h = \begin{pmatrix} D + M_z + \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z & \frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp^* & M_\perp \\ \frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp & 0 & \frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp^* \\ M_\perp^* & \frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp & D + M_z - \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z \end{pmatrix}. \tag{S2}$$

Case 1: Zero strain, zero off-axis magnetic field

For zero strain/electric field ($\{M_x, M_y, M_z\} = 0$) and zero off-axis magnetic field ($B_\perp = 0$), the Hamiltonian in Eqn. S2 is diagonal:

$$\mathbf{H}_0/h = \begin{pmatrix} D + \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D - \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z \end{pmatrix}, \tag{S3}$$

and the energy levels are given by the zero-field splitting D and Zeeman energies $\pm\frac{\gamma_{\mathrm{NV}}}{2\pi}B_z$,

$$E_{|\pm1,0\rangle}/h = \{D \pm \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z, 0\}, \tag{S4}$$

where $|\pm 1, 0\rangle$ are the Zeeman eigenstates

$$|+1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, |-1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{and } |0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{S5}$$

NV spin ensemble measurements in the DQ basis, for which the difference between the $f_{-1} = E_{|0\rangle\rightarrow|-1>}$ and $f_{+1} = E_{|0\rangle\rightarrow|+1>}$ transitions is probed (see Fig. 1b), are to first-order insensitive to inhomogeneities and fluctuations in D (e.g., due to drift in temperature), and other common-mode noise sources. However, DQ measurements are twice as sensitive to magnetic fields along B_z. The DQ basis therefore provides both enhanced magnetic field sensitivity and protection against common-mode noise sources (for higher order effects see, e.g., the Supplement of Ref. [S17]).

Case 2: Non-zero strain, zero off-axis magnetic field

For non-zero strain/electric field components, but negligible off-axis magnetic fields ($B_\perp \approx 0$), the energy eigenvalues of the NV Hamiltonian (Eqn. S2) for the $|\pm 1\rangle$ states become

$$E_{|\pm1\rangle}/h = D + M_z \pm \sqrt{(\frac{\gamma_{\mathrm{NV}}}{2\pi}B_z)^2 + ||M_\perp||^2} \tag{S6}$$

$$\approx D + M_z \pm \left[\frac{\gamma_{\mathrm{NV}}}{2\pi}B_z + \frac{||M_\perp||^2}{2\frac{\gamma_{\mathrm{NV}}}{2\pi}B_z} + \mathcal{O}\left(\frac{(||M_\perp||^4}{B_z^2}\right)\right]. \tag{S7}$$

From Eqn. S7 it follows that off-axis strain ($\propto ||M_\perp||$) is suppressed by moderate on-axis bias fields by a factor $\frac{||M_\perp||}{\gamma_{\mathrm{NV}}B_z/\pi}$, as noted in the main text. Reported values for $||M_\perp||$ are $\sim 10\,\mathrm{kHz}$ [S17] and $\sim 100\,\mathrm{kHz}$ [S18] for single NV centers in bulk diamond, and $\sim 7\,\mathrm{MHz}$ in nano-diamonds [S18]. Fig. 1c in the main text shows that the measured on-axis strain M_z in Sample B varies by $2-3\,\mathrm{MHz}$ (see Suppl. IV for details).

Case 3: Non-zero off-axis magnetic field

For non-zero off-axis magnetic field ($B_\perp \neq 0$) we find the energy values for the NV Hamiltonian (Eqn. S1) by treating $B_\perp$ as a small perturbation, with perturbation Hamiltonian $\mathbf{V} \equiv \mathbf{H} - \mathbf{H}_0$. To simplify the analysis we set $M_{||} = M_\perp = 0$. Using time-independent perturbation theory (TIPT, see for example Ref. [S19]), the corrected energy levels are then given by $E_{|\pm1,0\rangle} \approx E^{(0)}_{|\pm1,0\rangle} + E^{(1)}_{|\pm1,0\rangle} + E^{(2)}_{|\pm1,0\rangle} + \dots$, where $E^{(0)}_{|\pm1,0\rangle}$ are the bare Zeeman energies as given in Eqn. S4 and $E^{(k)}_{|\pm1,0\rangle}$ for $k > 0$ are the k-th order corrections. The energy corrections at first and second order are:

$$E^{(1)}_{|\pm1,0\rangle} = \langle \pm1, 0|\mathbf{V}|\pm1, 0\rangle = 0, \tag{S8}$$

and

$$E^{(2)}_{|\pm1\rangle} = \frac{\|\langle\mp1|\mathbf{V}|\pm1\rangle\|^2}{E_{|\pm1\rangle} - E_{|\mp1\rangle}} + \frac{\|\langle0|\mathbf{V}|\pm1\rangle\|^2}{E_{|\pm1\rangle}} = \frac{\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D \pm \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z} \approx \frac{\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D}, \tag{S9}$$

$$E^{(2)}_{|0\rangle} = \frac{\|\langle+1|\mathbf{V}|0\rangle\|^2}{-E_{|+1\rangle}} + \frac{\|\langle-1|\mathbf{V}|0\rangle\|^2}{-E_{|-1\rangle}} = -\left(\frac{\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D + \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z} + \frac{\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D - \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z}\right) \approx -\frac{2\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D}, \tag{S10}$$

where we have used in the last two lines the fact that $\frac{\gamma_{\mathrm{NV}}}{2\pi}B_z \ll D$ in our experiments. The new transition frequencies for $E_{|0\rangle\rightarrow|\pm1\rangle}$ are then found to be

$$f_{\pm1} \approx D + \frac{3\|\frac{\gamma_{\mathrm{NV}}}{2\pi}B_\perp\|^2}{D} \pm \frac{\gamma_{\mathrm{NV}}}{2\pi}B_z. \tag{S11}$$

From Eqn. S11 it follows that energy level shifts due to perpendicular magnetic fields are mitigated by the large zero-field splitting D; and are further suppressed in the DQ basis, as they add (approximately) in common-mode. At moderate bias fields, $B_z = 2-20\,\mathrm{mT}$, and typical misalignment angles of $\theta \sim 3^\circ$ (or lower), we estimate a frequency shift of $0.1-1\,\mathrm{kHz}$ in the SQ basis.

VIII. SPIN BATH DRIVING MODEL

The effective magnetic field produced by the ensemble of nitrogen spins is modeled as a Lorentzian line shape with spectral width δ_{N} (half width at half max) and a maximum $\gamma_{\mathrm{NV\text{-}N}}$ at zero drive frequency ($\Omega_{\mathrm{N}} = 0$). This lineshape is

derived in the context of dilute dipolar-coupled spin ensembles using the methods of moments [S15, Ch. III and IV] and is consistent with NV DEER linewidth measurements (see Suppl. XI). The limit to the NV ensemble T_2^* taking the bath drive into account is given by (see Eqn. 5 of main text)

$$1/T^*_{2,\text{NV-N}}(\Omega_\text{N}) = \Delta m \times \gamma_\text{NV-N} \frac{\delta_\text{N}^2}{\delta_\text{N}^2 + \Omega_\text{N}^2}. \tag{S12}$$

At sufficiently high drive strengths ($\Omega_\text{N} \gg \delta_\text{N}$), the nitrogen spin ensemble is coherently driven and the resulting magnetic field noise spectrum is detuned away from the zero-frequency component, to which NV Ramsey measurements are maximally sensitive [S20]. For this case, the NV spin ensemble T_2^* increases $\propto \Omega_\text{N}^2/\delta_\text{N}^2$. At drive strength $\Omega_\text{N} \lesssim \delta_\text{N}$, however, the nitrogen spin ensemble is inhomogeneously driven and the dynamics of the spin bath cannot be described by coherent driving. Nonetheless, $1/T_2^*$ given by Eqn. S12 approaches $\gamma_\text{NV-N}$ in the limit $\Omega_\text{N} \to 0$, which is captured by the Lorentzian model.

This model (Eqn. S12) is in excellent agreement with the data for Sample B ([N] = 0.75 ppm, $\delta_\text{N} \approx 11$ kHz), for which $\Omega_\text{N} > \delta_\text{N}$ for the range of drive strengths employed. $\Omega_\text{N} > \delta_\text{N}$ also holds when the slight mismatch of nitrogen spin resonances is taken into account, effectively increasing the nitrogen linewidth relevant for bath driving ($\delta_\text{N} \approx 60$ kHz, see discussion in main text). For Sample C ([N] = 10 ppm, $\delta_\text{N} \approx 150$ kHz), we find that the effective linewidth δ_N extracted from fitting the data in Fig. 4b is about 4× larger (≈ 600 kHz) than what is expected from the dipolar estimate even after account for the small B_0 misalignment angle and resultant slight mismatch of nitrogen spin resonance frequencies. We attribute this discrepancy to incoherent dynamics at drive strength $\Omega_\text{N} \sim \delta_\text{N}$. Indeed, we find that for Sample C at drive strengths $\Omega_\text{N} \lesssim \delta_\text{N}$ the Ramsey signals exhibit multi-exponential decay with slow and fast decay rates, consistent with a larger effective δ_N. To nonetheless enable a qualitative comparison with Sample B, in these instances the stretched exponential parameter is restricted to $p \geq 1$ when extracting the NV spin ensemble T_2^*. At drive frequencies $\Omega_\text{N} > \delta_\text{N}$, the observed Ramsey signal returns to a simple exponential decay, confirming the validity of our driving model in this regime for Sample C. A more complete driving model, beyond the scope of this work, should take into account the changes of spin bath dynamics at drive strengths $\Omega_\text{N} \sim \delta_\text{N}$.

IX. ^{14}N AND ^{15}N DOUBLE ELECTRON-ELECTRON RESONANCE SPECTRA

We account for the ^{14}N and ^{15}N spin resonances, observed in NV double electron-electron resonance (DEER) spectra (see Fig. 1d and S6), in terms of Jahn-Teller, hyperfine, and quadrupolar splittings. The relevant spin Hamiltonian for the substitutional nitrogen defect is given by [S21–S24]

$$H_\text{P1}/h = \mu_B/h\ \mathbf{B}\cdot\mathbf{g}\cdot\mathbf{S} + \mu_\text{N}/h\ \mathbf{B}\cdot\mathbf{I} + \mathbf{S}\cdot\mathbf{A}\cdot\mathbf{I} + \mathbf{I}\cdot\mathbf{Q}\cdot\mathbf{I} \tag{S13}$$

where μ_B is the Bohr magneton, h is the Planck's constant, $\mathbf{B} = (B_x, B_y, B_z)$ is the magnetic field vector, $\mathbf{g}$ is the electronic g-factor tensor, μ_N is the nuclear magneton, $\mathbf{S} = (\mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z)$ is the electronic spin vector, $\mathbf{A}$ is the hyperfine tensor, $\mathbf{I} = (\mathbf{I}_x, \mathbf{I}_y, \mathbf{I}_z)$ is the nuclear spin vector, and $\mathbf{Q}$ is the nuclear electric quadrupole tensor. This Hamiltonian can be simplified in the following way: First, we neglect the nuclear Zeeman energy (second term above) since its contribution is negligible at magnetic fields used in this work ($\simeq 10$ mT). Second, the Jahn-Teller distortion defines a symmetry axis for the nitrogen defect along any of the [111]-crystal axis directions [S25, S26]. Under this trigonal symmetry (as with NV centers), and by going into an appropriate coordinate system, tensors $\mathbf{g}$, $\mathbf{A}$, and $\mathbf{Q}$ are diagonal and defined by at most two parameters:

$$\mathbf{g} = \begin{pmatrix} g_\perp & 0 & 0 \\ 0 & g_\perp & 0 \\ 0 & 0 & g_\parallel \end{pmatrix}, \mathbf{A} = \begin{pmatrix} A_\perp & 0 & 0 \\ 0 & A_\perp & 0 \\ 0 & 0 & A_\parallel \end{pmatrix}, \text{ and } \mathbf{Q} = \begin{pmatrix} P_\perp & 0 & 0 \\ 0 & P_\perp & 0 \\ 0 & 0 & P_\parallel \end{pmatrix}. \tag{S14}$$

Here, $g_\perp$, $g_\parallel$, $A_\perp$, $A_\parallel$, $P_\perp$, and $P_\parallel$ are the gyromagnetic, hyperfine, and quadrupolar on- and off-axis tensor components, respectively, in the principal coordinate system. Further simplifications can be made by noting that the g-factor is approximately isotropic [S21], i.e., $g_\perp \approx g_\parallel \equiv g$, and that for exact axial symmetry the off-axis components of the quadrupole tensor, $P_\perp$, vanish [S27]. Equation S13 may now be written as

$$H_\text{P1}/h = g\mu_B/h\ B_z\ \mathbf{S}_z + A_\parallel \mathbf{S}_z\cdot\mathbf{I}_z + A_\perp(\mathbf{S}_x\cdot\mathbf{I}_x + \mathbf{S}_y\cdot\mathbf{I}_y) + P_{||}(\mathbf{I}_z^2 - \mathbf{I}^3/3). \tag{S15}$$

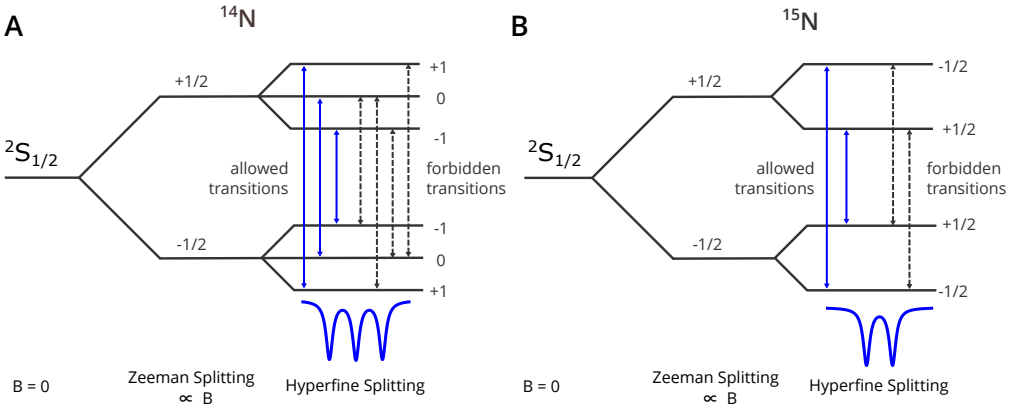


Figure S5. ^{14}N and ^{15}N spin energy level diagram.

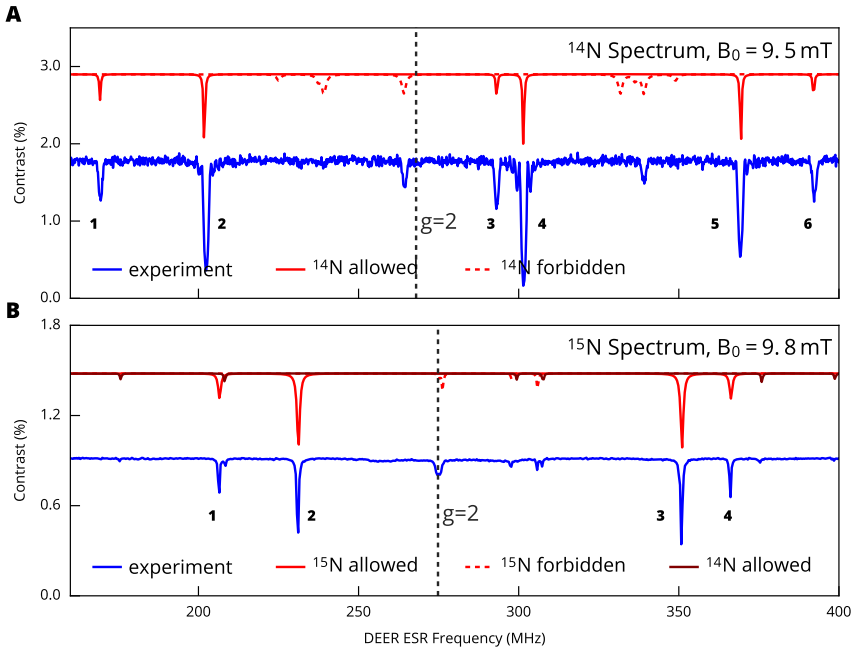


Figure S6. ^{14}N and ^{15}N DEER spectra. (a) Simulated (red) and measured (blue) ^{14}N DEER spectra for Sample B ([N] = 0.75 ppm, ^{14}N). Dipole-allowed nitrogen hyperfine transitions are labeled 1 - 6. Smaller peaks are attributed to degenerate forbidden hyperfine transitions ($\Delta m_I \neq 0$) of the off-axis nitrogen orientations. Frequencies are simulated using Eqn. S15 and plotted as Lorentzians with widths and amplitudes chosen to reflect the experimental data. Allowed hyperfine transitions have an approximate amplitude ratio of 1:3:1:3:3:1 (see main text). The Larmor frequency of an electronic spin without hyperfine shift ($g = 2$) is indicated as dashed black line. (b) Simulated (red) and measured (blue) ^{15}N DEER spectrum for Sample C ([N] = 10 ppm, ^{15}N). Dipole-allowed hyperfine transitions are labeled 1 - 4. Smaller peaks are attributed to forbidden ^{15}N hyperfine transitions and $g = 2$ dark spins. The spectrum of a small abundance of ^{14}N spins ($\approx 5\%$ of ^{15}N density) is visible as well.

^{14}N spectrum

^{14}N has $S = 1/2$ and $I = 1$, leading to six eigenstates $|m_S = \pm 1/2, m_I = 0, \pm 1\rangle$. The corresponding three dipole-allowed transitions ($\Delta m_S = \pm 1, \Delta m_I = 0$, solid arrows) are shown in Fig. S5, along with the four first-order forbidden transitions ($\Delta m_S = \pm 1, \Delta m_I = \pm 1$, dashed arrows). A nitrogen defect in diamond undergoes a Jahn-Teller (JT) distortion, which defines a hyperfine quantization axis along any of the four [111] crystallographic directions, irrespective of the applied magnetic field. Taking all JT orientations into account, the full ^{14}N spin resonance spectrum displays a total of 12 dipole-allowed resonances. By aligning the magnetic field along any of the

[111]-directions of the diamond crystal, the 12 transitions are partially degenerate and reduce to six visible transitions in an NV DEER measurement, with an amplitude ratio 1:3:1:3:3:1, as shown in Fig. 2b of the main text and Fig. S6a. We obtain the spectrum for the off-axis and degenerate JT orientations from Eqn. S15 by rotating the bias field by $\theta = 109.471$ around either the x or y axis, where θ is the angle between any two crystallographic axes, i.e., taking $\mathbf{B} \rightarrow \mathbf{R}_{\text{x or y}}(\theta = 109.471°) \cdot \mathbf{B}$.

In Fig. S6a, the simulated (degenerate) spectrum for ^{14}N spins is shown together with experimental data from Sample B ([N] = 0.75 ppm). A magnetic field $B_z = 95.5$ mT is applied along one of the [111]-orientations. For the simulation the following parameters have been used: $g\mu_B/h \approx 2.8025 \times 10^4$ MHz/T, where $g = 2.0025$ is the P1 electronic g-factor [S21], $\mu_B = 9.274 \times 10^{-24}$ J/T is the Bohr magneton, $h = 6.626 \times 10^{-34}$ Js is Planck's constant, $A_{||} = 114$ MHz, $A_\perp = 81.3$ MHz [S21, S22, S24], and $P_{||} = -3.97$ MHz [S22].

	^{14}N
g	2.0025 [S21]
$A_\perp, A_{||}$	114 MHz, 81.3 MHz [S21, S22, S24]
$P_{||}$	-3.97 MHz [S22]
	^{15}N
$A_{||}$, $A_\perp$	-159.7 MHz, -113.83 MHz [S24]
P	0 (since $I < 1$)

Table S2. Summary of defect parameters used to simulate the nitrogen resonance spectrum using Eqn. S15.

^{15}N spectrum

^{15}N has $S = 1/2$ and $I = 1/2$, leading to the four eigenstates $|m_S = \pm 1/2, m_I = \pm 1/2\rangle$. The corresponding two dipole-allowed transitions ($\Delta m_S = \pm 1, \Delta m_I = 0$, solid arrows) are shown in Fig. S5b, along with the two first-order forbidden transitions ($\Delta m_S = \pm 1, \Delta m_I = \pm 1$, dashed arrows). The experimental NV DEER spectrum for Sample C ([^{15}N]=10 ppm) is shown in Fig. S6b, along with a simulated ^{15}N spectrum. For the ^{15}N simulation we used $B_0 = 9.8$ mT, $A_{||} = -159.7$ MHz, $A_\perp = -113.83$ MHz [S24], and $P_{||} = 0$ (since $I < 1$).

X. CONTINUOUS VERSUS PULSED SPIN BATH DRIVING

As described in the main text, both continuous (CW) and pulsed driving can decouple the electronic spin bath from the NV sensor spins (see Fig. S7). In CW driving, the bath spins are driven continuously such that they undergo many Rabi oscillations during the characteristic interaction time $1/\gamma_{\text{NV-N}}$, and thus the time-averaged NV-N dipolar interaction approaches zero. For pulsed driving, π-pulses resonant with spin transitions in the bath are applied midway through the NV Ramsey free precession interval, to refocus bath-induced dephasing. Fig. S7a illustrates both methods for a given applied RF field with a Rabi frequency of Ω_{N}.

Although we treat CW driving in the main text in detail, we find experimentally that pulsed driving yields similar T_2^* improvements over the measured range of Rabi drives. For example, Fig. S7a compares T_2^* for Sample B for both schemes at maximum bath drive strength $\Omega_{\text{N}} = 1.5$ MHz (for pulsed driving $\tau_\pi \equiv 1/2\Omega_{\text{N}}$). Both decoupling schemes result in comparable T_2^* improvements ($13 - 15\times$) over the non-driven SQ measurement, which is shown for reference. We attribute the slightly lower max T_2^* achieved in pulsed driving to detunings of the RF drive from the spin resonances of the main nitrogen groups, leading to less efficient driving of the spin population (see next section). To study the efficacy of both driving schemes, we plot T_2^* as a function of Rabi drive Ω_{N} in Fig. S7b. In the limit of $\tau_\pi \approx T_2^*$, pulsed driving resembles the CW case and both schemes converge to the same maximal T_2^*.

Despite the similar improvements in T_2^* achieved using both methods, pulsed driving can reduce heating of the MW delivery loop and diamond sample - an important consideration for temperature sensitive applications. For this reason, pulsed driving may be preferable in such experiments despite the need for π-pulse calibration across multiple resonances.

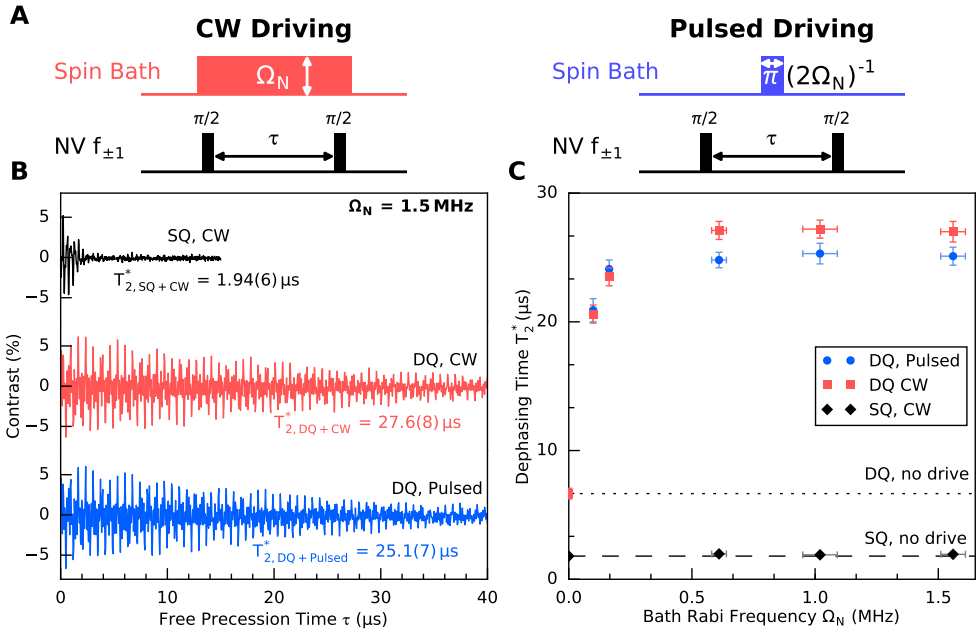


Figure S7. Comparison of CW to pulsed spin bath driving experiments using Sample B. (a) Sequences used for CW (left) and pulsed (right) decoupling of the electronic nitrogen spin bath. For both methods, six distinct frequencies are used to resonantly address the nitrogen spin bath via an applied RF field with equal Rabi frequency Ω_N on each spin transition. In (b) the Ramsey decay in the DQ basis for CW and pulsed driving are compared ($\Omega_N = 1.5$ MHz). The decay for CW driving in the SQ basis is included for reference. (c) Depicts T_2^* as a function of bath Rabi frequency for DQ CW (red squares) and DQ pulsed (blue circles) spin bath driving, with the SQ CW results (black diamonds) again included for reference. The finely (coarsely) dashed line indicates the T_2^* value in the DQ (SQ) basis without any drive field applied to the bath spins.

XI. NV AND NITROGEN SPIN RESONANCE LINEWIDTH MEASUREMENTS

The NV and nitrogen (P1) ensemble spin resonance linewidths are determined using pulsed ESR and pulsed DEER NV spectral measurements, respectively, as shown in Fig. S8. Low Rabi drive strength and consequently long π-pulse durations can be used to avoid Fourier power broadening [S28]. We find that nitrogen spin resonance spectra are typically narrower than for NV ensembles in the SQ basis, due to the effects of strain gradients in diamond on NV zero-field splittings.

For the spin bath driving model described in the main text (Eqn. 4), we are interested in the natural (i.e., non-power-boadened) linewidth δ_N of spin resonances corresponding to, for example, ^{14}N groups $1-6$ (see Fig. 2b in main text and Fig. S8a in Supplement). In Ref. [S29] it was reported that the different ^{14}N groups have approximately equal linewidth, i.e., that $\delta_{N,i} \approx \delta_N$. However, we find that the bias field B_z being only slightly misaligned (~3 degree) from one of the [111] crystal axes causes the three degenerate spin resonances to be imperfectly overlapped, leading to a larger effective linewidth.

In Fig. S8b and c we compare the NV pulsed DEER linewidths of ^{14}N group 1 (a single resonance) with that of group 5 (three overlapped resonances) for different π_{bath}-pulse durations. At short π_{bath}-pulse durations (high MW powers), the linewidths are power broadened due to the applied microwave field, such that the measured linewidth is a convolution of the natural linewidth and the inverse duration of the π_{bath}-pulse [S28]. At longer π_{bath}-pulse durations (reduced MW power), however, the measured linewidth approaches its natural width. In this instance, and for dipolar-limited linewidth broadening, the lineshape is Lorentzian with full width at half max $\Gamma = 1/\pi T^*_{2,N}$. At the longest π_{bath}-pulse durations used in this work, we find that group 1 consists of a narrow, approximately 25 kHz-wide peak. In contrast, group 5 reveals two peaks, consisting of two overlapped ^{14}N transitions and one detuned transition, which is attributed to imperfect magnetic field alignment. The splitting between the two peaks in group 5 is $\approx$ 80 kHz, which we use as the effective ^{14}N linewidth δ_N in Eqn. 4 of the main text, and which is consistent with the value extracted from fitting the spin-bath driving model to the data (see Fig. 4a, $\delta_N^{fit} \approx 60$ kHz).

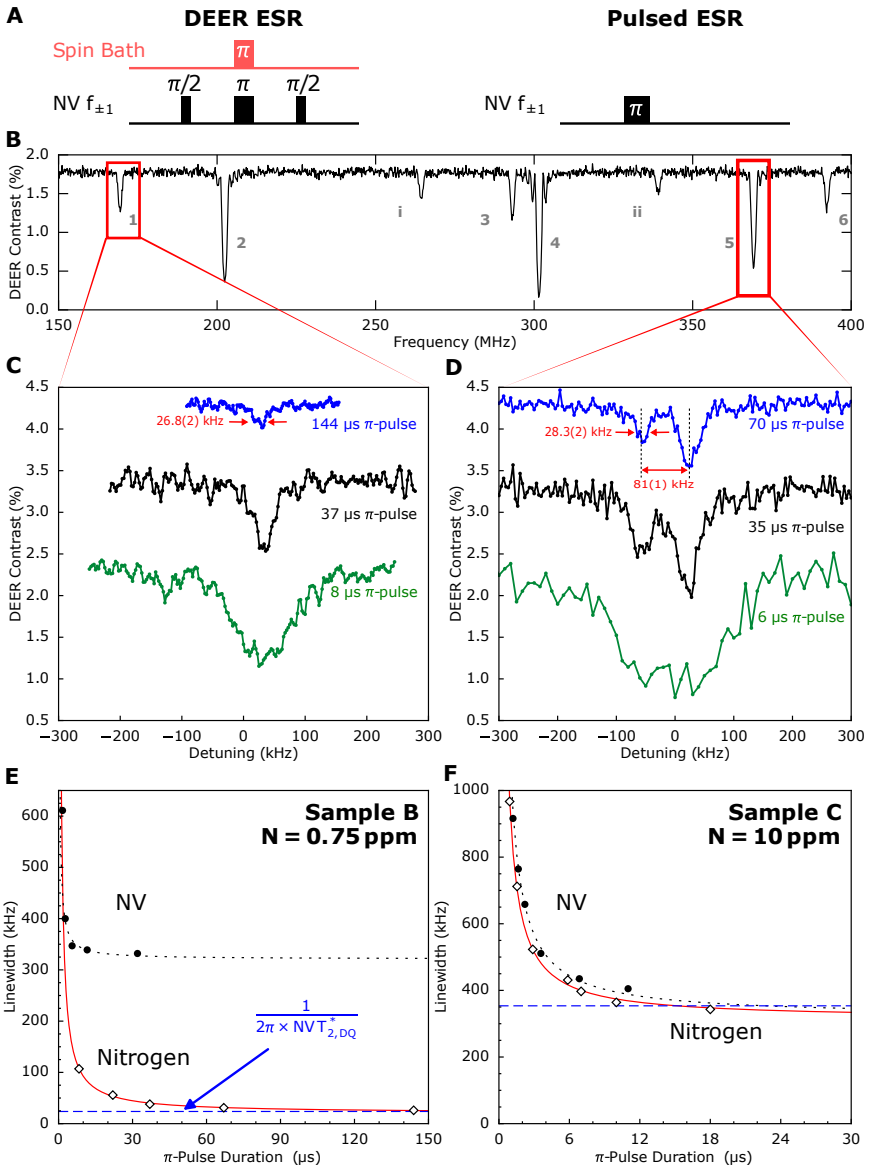


Figure S8. Comparison of nitrogen and NV spin resonance linewidths. (a) The pulsed DEER (left) sequence used for spin resonance measurements of the nitrogen spins and the pulsed ESR (right) sequence is used for NV SQ spin resonance measurements. (b) DEER spectrum from 150 to 400 MHz including all six nitrogen transitions and two forbidden transitions in Sample B. (c) DEER spectra of a single nitrogen transition are shown for three different bath π-pulse durations. A minimum measured linewidth of 26.8(2) kHz was recorded using a 144 µs π-pulse. (d) DEER spectra for a group containing three nearly degenerate off-axis nitrogen transitions. When bath π-pulses of 70 µs and 35 µs are used, two features are resolved corresponding to a single nitrogen transition detuned by 81 kHz from two nearly overlapped transitions. (e) Comparison of the NV ESR linewidth (black dots) and the DEER linewidth for a single nitrogen transition (diamonds) as a function of π-pulse duration for Sample B ([N] = 0.75 ppm). The fine, black dashed and red solid lines correspond to fits of the NV and nitrogen spin resonance linewidths to the functional form $a/x + b$, where b is the saturation linewidth. The coarse, blue dashed line indicates the expected linewidth from the measured NV T_2^* in the DQ basis (assuming a Lorentzian linewidth). (f) Same as (e) but for Sample C ([N] = 10 ppm).

In Fig. S8e we compare the measured NV and ^{14}N group 1 ensemble linewidths (full width at half max) for

Sample B as a function of π-pulse duration. For both species, the linewidth narrows at long π-pulse durations, as discussed above, reaching non-power-broadened (natural) values. Notably, the non-power-broadened NV linewidth [321(7) kHz, extracted from a fit to the data] is $\sim 16\times$ larger than the natural ^{14}N linewidth [20.6(1.2) kHz]. This order-of-magnitude difference is a manifestation of the strong strain field gradients in this sample. Specifically, pulsed ESR measurements of the NV ensemble linewidth (see Fig. S8a) are performed in the SQ $\{0,+1\}$ or $\{0,-1\}$ sub-basis, and are therefore strain gradient limited. In contrast, nitrogen defects in diamond have $S = 1/2$, and thus do not couple to electric fields or strain gradients. As a consistency check, note that NV ensemble Ramsey measurements in Sample B, made in the DQ basis (with no spin-bath driving), yield a strain-independent dephasing time $T^*_{2,\mathrm{DQ}} = 6.9(5)\,\mu$s. This dephasing time, presumably limited by the nitrogen spin bath, implies a ^{14}N spin resonance linewidth given by $\frac{1}{2} \times 1/\pi T^*_{2,\mathrm{DQ}} = 23(2)$ kHz, which is in good agreement with our pulsed DEER measurements of the natural ^{14}N linewidth. Similar consistency is found for measurements of the NV and ^{15}N ensemble spin resonance linewidths in Sample C, as shown in Fig. S8f. Such agreement across multiple samples is further evidence that the DQ T_2^* value for NV ensembles is limited by the surrounding nitrogen spin bath, as discussed in the main text. Note that for our samples [NV] $\ll$ [N] and we can therefore ignore the back action of NVs onto nitrogen spins in the DEER readout. For denser NV samples, however, this back action has to be taken into account [S30].

XII. DC MAGNETOMETRY WITH DQ AND SPIN-BATH DRIVE

Assuming a signal-to-noise ratio of unity, the minimum detectable magnetic field δB_{min} in a Ramsey measurement is given by [S17]

$$\delta B_{min} \approx \frac{\delta S}{\max|\frac{\partial S}{\partial B}|}, \tag{S16}$$

where the Ramsey signal S is

$$S = C(\tau)\sin(\gamma_{\mathrm{NV}} B_{DC}\tau). \tag{S17}$$

Here, $C(\tau) = C_0 \exp\left(-(\tau/T_2^*)^p\right)$ is the time-dependent measurement contrast defined via the NV spin-state-dependent fluorescence visibility (see Suppl. I), γ_{NV} is the NV gyromagnetic ratio, B_{DC} is the magnetic field to be sensed, and τ is the sensing time during which the NV sensor spins accumulate phase. The term $\max|\frac{\partial S}{\partial B}|$ is the maximum slope of the Ramsey signal,

$$\max|\frac{\partial S}{\partial B}| = C(\tau)\gamma_{\mathrm{NV}}\tau. \tag{S18}$$

Assuming uncorrelated, Gaussian noise, $\delta S = \sigma(\tau)/\sqrt{n_{meas}}$ is the standard error of the contrast signal, which improves with number of measurements n_{meas}. Including a dead time τ_D that accounts for time spent during initialization of the NV ensemble and readout of the spin-state-dependent fluorescence during a single measurement, $n_{meas} = T/(\tau+\tau_D)$ measurements are made over the total measurement time T. δB_{min} is then found to be

$$\delta B_{min} = \frac{\sigma\sqrt{\tau+\tau_D}}{C(\tau)\gamma_{\mathrm{NV}}\tau\sqrt{T}}, \tag{S19}$$

and the sensitivity is given by multiplying δB_{min} by the bandwidth $\sqrt{T}$ and including a factor $\Delta m = 1(2)$ for the SQ (DQ) basis:

$$\eta = \frac{\delta B_{min}\sqrt{T}}{\Delta m} = \frac{\sigma\sqrt{\tau+\tau_D}}{\Delta m \times C(\tau)\gamma_{\mathrm{NV}}\tau}. \tag{S20}$$

Note that in the ideal case, $\tau_D \ll \tau$, we have $\frac{\sqrt{\tau+\tau_D}}{\tau} \approx 1/\sqrt{\tau}$ and the sensitivity η scales $\propto \tau^{-1/2}\exp\left(\tau/T_2^*\right)^p$. The optimal sensing time in our Ramsey experiment is then $\tau_{opt} \approx T_2^*/2$ for $p = 1-2$. However, in the more realistic case, $\tau_D \sim \tau$, the improvement of η with increasing τ approaches a linear scaling and $\eta \propto \tau^{-1}\exp\left(\tau/T_2^*\right)^p$ for $\tau_D \gg \tau$. The optimal sensing time then becomes $\tau_{opt} \approx T_2^*$. Consequently, the measured increase in sensitivity may exceed the enhancement estimated from the idealized case without overhead time.

With Eqn. S20 we calculate and compare the sensitivities for the three measurement modalities (SQ, DQ, and DQ + spin-bath drive) applied to Sample B. Using $C \approx 0.026$, which remains constant for the three schemes (see Fig. S9a), sensing times $\tau_{SQ} = 1.308\,\mu s$, $\tau_{DQ} = 6.436\,\mu s$, and $\tau_{DQ+Drive} = 23.99\,\mu s$, standard deviations $\sigma_{SQ} = 0.0321$, $\sigma_{DQ} = 0.0324$, and $\sigma_{DQ+Drive} = 0.0325$ calculated from 1 s of data, fixed sequence duration of $\tau + \tau_D = 70\,\mu s$, and $\gamma_{NV} = 2\pi \times 28\,\mathrm{GHz/T}$, the estimated sensitivities for the SQ, DQ and DQ+Drive measurement schemes are $\eta = 70.7$, 6.65, and $1.97\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$, respectively. In summary, we obtain a $10\times$ improvement in DC magnetic field sensitivity in the DQ basis, relative to the conventional SQ basis, and a $35\times$ improvement using the DQ basis with spin bath drive. Note that this enhancement greatly exceeds the expected improvement when no dead time is present ($\tau_D \ll \tau$) and is attributed to the approximately linear increase in sensitivity with sensing times $\tau \lesssim \tau_D$. We also plot the Allan deviation for the three schemes in Fig. S9b showing a $\tau^{-1/2}$ scaling for a measurement time of $\approx 1\,s$ and the indicated enhancements in sensitivity.

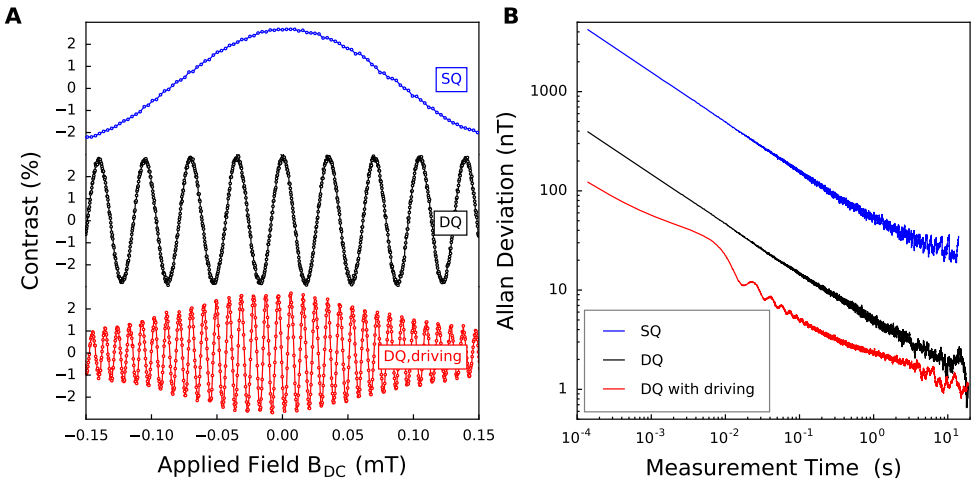


Figure S9. DC magnetic field sensing and Allan deviation. (a) DC magnetometry curves for SQ, DQ, and DQ with spin-bath driving in Sample B, produced by sweeping the magnitude of a coil-generated applied magnetic field (in addition to the fixed bias field) while the free precession interval τ is set to $\tau_{SQ} = 1.308\,\mu s$ (blue, top), $\tau_{DQ} = 6.436\,\mu s$ (black, middle), and $\tau_{DQ,drive} = 23.990\,\mu s$ (red, bottom). Reproduced from Fig. 4d (b) Allan deviation using the same fixed τ values from (a) for measurements using SQ (blue), DQ (black), and DQ with driving (red). The external field strength was tuned to sit on a zero crossing of the respective DC magnetometry curves in (a) for sine magnetometry.

Lastly, we discuss an appropriate choice of spin concentrations in diamond and other sample material properties for enhanced-sensitivity magnetometry employing DQ coherence and spin bath driving. To simplify the discussion, we focus on the following combination of relevant parameters for NV magnetometry, which (in the appropriate limits discussed herein) is proportional to the photon-shot-noise-limited volume-normalized magnetic sensitivity η^{V} (see Suppl. of Ref. [S7]), given by

$$\eta^{\mathrm{V}} \propto \left[\Delta m \times \sqrt{n_{\mathrm{NV}} \cdot [\mathrm{N}] \cdot T_2^*([\mathrm{N}])}\right]^{-1} \equiv \eta_{\mathrm{N}}. \tag{S21}$$

Here, $\Delta m = 1(2)$ in the SQ (DQ) basis, [N] is the substitutional nitrogen (P1) center concentration, $n_{\mathrm{NV}} = [\mathrm{NV}]/[\mathrm{N}]$ is the normalized concentration of NV centers relative to the nitrogen concentration, and $T_2^*([\mathrm{N}])$ is the NV ensemble dephasing time in the sensing basis chosen (SQ or DQ). The quantity η_{N} describes the dependence of the sensitivity on nitrogen concentration [N]. Since the nitrogen concentration enters Eqn. S21 both explicitly and also through n_{NV} and T_2^*, we need to investigate Eqn. S21 for a range of [N].

In the case of fixed n_{NV} and nitrogen-spin-bath-limited dephasing, i.e., $T_2^* \propto 1/[\mathrm{N}]$, η_{N} remains constant as a function of [N]. In this simplistic picture, shorter T_2^* values may be exchanged for higher nitrogen (and thus NV center) concentrations and vice versa, with no effect on sensitivity. Such a discussion, however, neglects experimental overhead due to NV state initialization and readout, which is characterized by the dead time τ_D (see Eqn. S20). Since $T_2^* \lesssim 1\,\mu s \lesssim \tau_D$ in a typical SQ NV ensemble experiment, increasing T_2^* through optimized sample fabrication, DQ coherence magnetometry, and/or spin bath driving is preferred, and larger sensitivity gains are obtained when compared to an equivalent increase in NV center concentration.

More generally, the ensemble T_2^* depends on numerous diamond-related parameters (including the concentration of spin impurities and strain fields) and external conditions (such as temperature fluctuations of the diamond sample

and magnetic field gradients due to the applied bias field). Focusing on the parameters intrinsic to diamond, the relevant contributions to T_2^* are (compare to Eqn. 1 in main text)

$$1/T_2^* \approx 1/T_2^*\{\text{NV-}^{13}\text{C}\} + 1/T_2^*\{\text{NV-N}\}(\Omega_\text{N}) + 1/T_2^*\{\text{NV-NV}\}(\text{N}) + 1/T_2^*\{\text{strain}\} + ..., \tag{S22}$$

where we added the term $1/T_2^*\{\text{NV-NV}\}$ to account for dephasing due to NV-NV dipolar interactions. This dephasing mechanism was neglected in the main text due to the low N-to-NV conversion efficiencies of Samples A, B, and C ($n_\text{NV} \ll 1$), but its contribution becomes relevant at increased conversion efficiencies intended for optimized diamond magnetometry. To model Eqn. S21 across a range of nitrogen concentrations, we now combine Eqns. S21 and S22 and include the dependence of $T_2^*\{\text{NV-N}\}(\Omega_\text{N})$ on bath drive strength Ω_N (see Eqn. 5 in the main text). We also anticipate optimized diamond samples to be isotopically engineered with $T_2^*\{^{13}\text{C} = 0.01\,\%\} \simeq 100\,\mu\text{s}$ (or longer) and to possess strain field gradients comparable to this work's samples ($T_2^*\{\text{strain}\} \simeq 5\,\mu\text{s}$). N-to-NV conversion efficiencies of up to 30 % have been reported for NV ensembles [S31] suggesting that $n_\text{NV} \simeq 0.4$ is feasible for an optimized diamond sample. The simulation results for η_N in this parameter regime are summarized in Fig. S10 for SQ (blue) and DQ coherence magnetometry (red) and plotted for spin bath drive strengths $\Omega_\text{N} = 0$ (solid), 1, and 10 MHz (dashed).

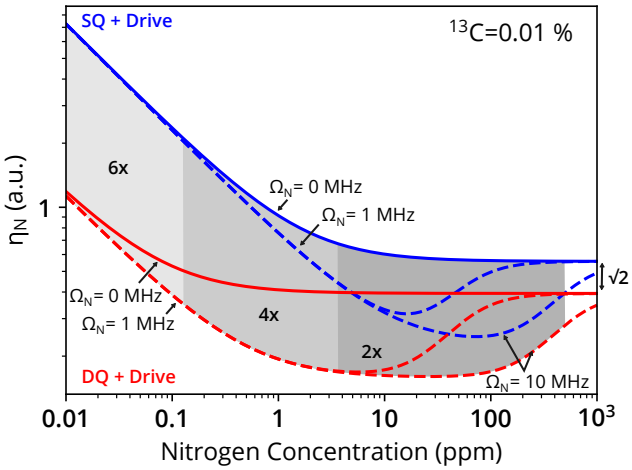


Figure S10. η_N given by Eqn. S21, which is proportional to the NV ensemble volume-normalized magnetic field sensitivity, as a function of nitrogen (P1) center concentration [N] for SQ (blue) and DQ (red) magnetometry with spin bath drive strengths $\Omega_\text{N} = 0$ (solid), 1, and 10 MHz (dashed). Lower values correspond to higher sensitivities and vice versa. For the simulation we combine Eqns. S21 and S22 with the following parameters: $T_2^*\{^{13}\text{C} = 0.01\,\%\} \simeq 100\,\mu\text{s}$, $T_2^*\{\text{strain}\} \simeq 5\,\mu\text{s}$, $T_2^*\{\text{NV-N}\}$ is given by Eqn. 5 of the main text, $n_\text{NV} = 0.4$, and $T_2^*\{\text{NV-NV}\}(\text{N}) = (A_\text{NV-NV} \cdot n_\text{NV} \cdot [\text{N}]/4)^{-1}$. Here, $A_\text{NV-NV} \approx 2A_\text{NV-N} \simeq 2\pi \times 33\,\text{kHz/ppm}$ due to the twice higher spin multiplicity of the NV centers [S15, Ch. III and IV], and the factor 1/4 accounts for the fraction of NV centers used for sensing when all four NV orientations are distinguished. In this instance and assuming perfect optical initialization of NV centers, 3/4 of the NV centers are in the $m_s = 0$ spin state and do not contribute to dephasing during the sensing sequence. Grey shaded regions indicate approximate improvements in sensitivity for DQ magnetometry with drive over SQ magnetometry alone.

Without spin bath drive applied ($\Omega_\text{N} = 0$) and at low nitrogen concentrations ($[\text{N}] \lesssim 1\,\text{ppm}$), η_N in Eqn. S21 is larger (i.e., sensitivity is reduced) for SQ magnetometry due to the T_2^*-limit imposed by strain gradients. In this lower nitrogen regime, working in the DQ basis leads to substantial improvements in sensitivity (i.e., smaller values and thus higher sensitivity). At higher nitrogen concentrations ($[\text{N}] \gtrsim 1\,\text{ppm}$), dipolar interactions with the spin bath dominate NV dephasing, strain contributions become negligible, and DQ coherence approaches a $\sqrt{2}$ enhancement in sensitivity over SQ coherence measurements. Note that the crossover between the low and high nitrogen regime is set by the strain contributions to T_2^* (here $\simeq 5\,\mu\text{s}$) and lower (higher) strain contributions shift the crossover to lower (higher) nitrogen concentrations.

With spin bath drive applied, however, we find that additional gains in sensitivity are obtained for both SQ and DQ coherence measurements. The largest improvements for η_N are obtained when DQ coherence measurements are employed at high spin bath drive strengths ($1 - 10\,\text{MHz}$) and for an optimal nitrogen regime of $1 - 100\,\text{ppm}$. In practice, we expect the $1 - 10\,\text{ppm}$ nitrogen regime to be optimal when all parameters relevant for magnetic sensing are considered. Herein we discuss additional parameters only qualitatively: i) The necessary Rabi frequency for effective spin bath driving increases linearly with nitrogen concentration, meaning the required RF power increases quadratically. Spin bath driving at high nitrogen concentrations becomes thus increasingly more challenging. ii)

Samples A, B, and C in the main text were selected to have a predominately electronic nitrogen (P1) spin bath. The incorporation of high nitrogen concentrations in diamond, however, can lead to a larger variety of nitrogen-related spin species in the bath, which include nitrogen-clusters, NV^0, and NVH defects (for example see Refs. [S32–S34] and therein). A more diverse spin bath severely increases the complexity of the bath drive, eventually rendering it impractical. Finally, iii) there are indications that diamond samples with a high nitrogen content exhibit a larger fraction of NV^0 centers. In such samples, the NV^- measurement contrast C (see Eqn. S20) and, thus magnetic sensitivity, is diminished.

Summarizing, our analysis suggests that the $1-10$ ppm nitrogen regime is optimal for high sensitivity magnetometry using NV ensembles but further work is required to quantitatively account for all parameters relevant for sensing. Note that the $2\times$, $4\times$, and $6\times$ enhancement in sensitivity indicated in Fig. S10 for DQ with drive over SQ magnetometry alone, corresponds to a $4\times$, $16\times$, and $36\times$ reduction in measurement time, respectively; and even larger relative enhancements in sensitivity should be realized when accounting for the experimental dead time τ_D in NV ensemble experiments.

XIII. DEPHASING CHANNELS PER SAMPLE

Table S3. NV spin ensemble dephasing mechanisms for Sample A. Individual contributions to dephasing are determined using the estimated/calibrated values described in the main text and Supplement (column 2). The data show good agreement between calculated and measured total dephasing times $T^*_{2,\mathrm{SQ}}$ and $T^*_{2,\mathrm{DQ}}$ (last two rows).

Channel	Magnitude	Dephasing		Method
		1/μs	μs	
strain	0.0028 MHz/μm	0.190	5	estimate
^{14}N	0.05 ppm	0.0029	348	dipolar estimate
^{13}C	0.01%	0.01	100	calibration
magnetic field gradient @ 20 G	0.000056 MHz/G	0.00112	893	estimate
total SQ		0.2035	4.9	$5-12$ μs (measured)
total DQ $\times 2$ (no strain)		0.014	71	68 μs (measured)

Table S4. NV spin ensemble dephasing mechanisms for Sample B. Similar to Sample A. Additionally, spin echo double-electron resonance (SEDOR) measurements were performed to estimate dephasing contributions from individual nitrogen resonance lines (for details see Ref. [S35]).

Channel	Magnitude	Dephasing		Method
		1/μs	μs	
strain	0.0028 MHz/μm	0.190	5	estimate
^{14}N (allowed)		0.056	18	SEDOR
^{14}N (forbidden)		0.0047	214	SEDOR
^{13}C	0.01%	0.01	100	calibration
magnetic field gradient @ 85 G	0.000056 MHz/G	0.00474	210	estimate
total SQ		0.265	3.8	$1-10$ μs (measured)
total DQ $\times 2$ (no strain)		0.076	13.1	13.8 μs (measured)

Table S5. NV spin ensemble dephasing mechanisms for Sample C, similar to Samples A and B.

Channel	Magnitude	Dephasing		Method
		1/μs	μs	
strain	0.0028 MHz/μm	0.140	7	estimate
^{15}N (allowed)		0.59	2	SEDOR
^{15}N (forbidden)		0.15	7	SEDOR
^{14}N (5% of N15)		0.0391	26	estimated
^{13}C	0.05%	0.05	20	calibrated
magnetic field gradient @ 100 G	0.000022 MHz/G	0.0022	446	estimate
total SQ		1.01	1.0	$0.3 - 1.2\,\mu$s (measured)
total DQ $\times 2$ (no strain)		0.87	1.1	1.2 μs (measured)

[S1] P Elleaume, O. Chubar, and J Chavanne, "Computing 3D magnetic fields from insertion devices," in *Proceedings of the 1997 Particle Accelerator Conference (Cat. No.97CH36167)*, Vol. 3 (IEEE, 1998) pp. 3509–3511.

[S2] Marcus W. Doherty, Neil B. Manson, Paul Delaney, Fedor Jelezko, Jörg Wrachtrup, and Lloyd C L Hollenberg, "The nitrogen-vacancy colour centre in diamond," Physics Reports **528**, 1–45 (2013).

[S3] N Bar-Gill, L.M. Pham, A Jarmola, D Budker, and R.L. Walsworth, "Solid-state electronic spin coherence time approaching one second," Nature Communications **4**, 1743 (2013).

[S4] Eisuke Abe, Alexei M. Tyryshkin, Shinichi Tojo, John J.L. Morton, Wayne M. Witzel, Akira Fujimoto, Joel W. Ager, Eugene E. Haller, Junichi Isoya, Stephen A. Lyon, Mike L.W. Thewalt, and Kohei M. Itoh, "Electron spin coherence of phosphorus donors in silicon: Effect of environmental nuclei," Physical Review B - Condensed Matter and Materials Physics **82**, 121201 (2010).

[S5] V. M. Acosta, E. Bauch, M. P. Ledbetter, C. Santori, K. M C Fu, P. E. Barclay, R. G. Beausoleil, H. Linget, J. F. Roch, F. Treussart, S. Chemerisov, W. Gawlik, and D. Budker, "Diamonds with a high density of nitrogen-vacancy centers for magnetometry applications," Physical Review B - Condensed Matter and Materials Physics **80**, 1–15 (2009).

[S6] V. M. Acosta, E. Bauch, M. P. Ledbetter, A. Waxman, L.-S. Bouchard, and D. Budker, "Temperature Dependence of the Nitrogen-Vacancy Magnetic Resonance in Diamond," Physical Review Letters **104**, 070801 (2010).

[S7] John F. Barry, Matthew J. Turner, Jennifer M. Schloss, David R. Glenn, Yuyu Song, Mikhail D. Lukin, Hongkun Park, and Ronald L. Walsworth, "Optical magnetic detection of single-neuron action potentials using quantum defects in diamond," Proceedings of the National Academy of Sciences **113**, 14133–14138 (2016).

[S8] C. Grezes, B. Julsgaard, Y. Kubo, W. L. Ma, M. Stern, A. Bienfait, K. Nakamura, J. Isoya, S. Onoda, T. Ohshima, V. Jacques, D. Vion, D. Esteve, R. B. Liu, K. Mølmer, and P. Bertet, "Storage and retrieval of microwave fields at the single-photon level in a spin ensemble," Physical Review A **92**, 020301 (2015).

[S9] Toyofumi Ishikawa, Kai-Mei C. Fu, Charles Santori, Victor M. Acosta, Raymond G. Beausoleil, Hideyuki Watanabe, Shinichi Shikata, and Kohei M. Itoh, "Optical and Spin Coherence Properties of Nitrogen-Vacancy Centers Placed in a 100 nm Thick Isotopically Purified Diamond Layer," Nano Letters **12**, 2083–2087 (2012).

[S10] Paul V Klimov, Abram L Falk, David J Christle, Viatcheslav V Dobrovitski, and David D Awschalom, "Quantum entanglement at ambient conditions in a macroscopic solid-state spin ensemble," Science Advances **1**, e1501015–e1501015 (2015).

[S11] William F. Koehl, Berk Diler, Samuel J. Whiteley, Alexandre Bourassa, N. T. Son, Erik Janzén, and David D. Awschalom, "Resonant optical spectroscopy and coherent control of Cr4+ spin ensembles in SiC and GaN," Physical Review B **95**, 035207 (2017).

[S12] Georg Kucsko, Soonwon Choi, Joonhee Choi, Peter C Maurer, Hengyun Zhou, Renate Landig, Hitoshi Sumiya, Shinobu Onoda, Junich Isoya, Fedor Jelezko, Eugene Demler, Norman Y Yao, and Mikhail D Lukin, "Critical thermalization of a disordered dipolar spin system in diamond," arXiv preprint (2016).

[S13] David R. Glenn, Roger R. Fu, Pauli Kehayias, D. Le Sage, Eduardo A. Lima, Benjamin P. Weiss, and Ronald L. Walsworth, "Micrometer-scale magnetic imaging of geological samples using a quantum diamond microscope," Geochemistry, Geophysics, Geosystems **18**, 3254–3267 (2017).

[S14] M. J R Hoch and E. C. Reynhardt, "Nuclear spin-lattice relaxation of dilute spins in semiconducting diamond," Physical Review B **37**, 9222–9226 (1988).

[S15] A. Abragam, *The principles of nuclear magnetism* (Clarendon Press, 1983) p. 599.

[S16] Erik Bauch, Paul Jungyun, Connor Hart, Swati Singh, Jennifer M. Schloss, Matthew J. Turner, John F. Barry, Linh Pham, Nir Bar-Gill, Susanne F. Yelin, and Ronald Walsworth, "Quantum coherence of NV center solid-state spins in diamond," Manuscript in Preparation (2018).

[S17] Kejie Fang, Victor M. Acosta, Charles Santori, Zhihong Huang, Kohei M. Itoh, Hideyuki Watanabe, Shinichi Shikata, and Raymond G. Beausoleil, "High-Sensitivity Magnetometry Based on Quantum Beats in Diamond Nitrogen-Vacancy Centers," Physical Review Letters **110**, 130802 (2013).

[S18] P. Jamonneau, M. Lesik, J. P. Tetienne, I. Alvizu, L. Mayer, A. Dréau, S. Kosen, J.-F. Roch, S. Pezzagna, J. Meijer, T. Teraji, Y. Kubo, P. Bertet, J. R. Maze, and V. Jacques, "Competition between electric field and magnetic field noise in the decoherence of a single spin in diamond," Physical Review B **93**, 024305 (2016).

[S19] Jun John Sakurai and Jim Napolitano, *Modern quantum mechanics* (Dorling Kindersley, 2014).

[S20] Lukasz Cywiński, Roman M. Lutchyn, Cody P. Nave, and S. Das Sarma, "How to enhance dephasing time in superconducting qubits," Physical Review B **77**, 174509 (2008).

[S21] W. V. Smith, P. P. Sorokin, I. L. Gelles, and G. J. Lasher, "Electron-spin resonance of nitrogen donors in diamond," Physical Review **115**, 1546–1552 (1959).

[S22] R. J. Cook and D. H. Whiffen, "Electron Nuclear Double Resonance Study of a Nitrogen Centre in Diamond," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **295**, 99–106 (1966).

[S23] J H N Loubser and J A van Wyk, "Electron spin resonance in the study of diamond," Reports on Progress in Physics **41**, 1201–1248 (1978).

[S24] A Cox, M E Newton, and J M Baker, "13 C, 14 N and 15 N ENDOR measurements on the single substitutional nitrogen centre (P1) in diamond," Journal of Physics: Condensed Matter **6**, 551–563 (1994).

[S25] Gordon Davies, "The Jahn-Teller effect and vibronic coupling at deep levels in diamond," Reports on Progress in Physics **44**, 787–830 (1981).

[S26] C A J Ammerlaan and E. A. Burgemeister, "Reorientation of Nitrogen in Type-Ib Diamond by Thermal Excitation and Tunneling," Physical Review Letters **47**, 954–957 (1981).
[S27] Charles P. Slichter, *Principles of Magnetic Resonance*, Springer Series in Solid-State Sciences, Vol. 1 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1990) pp. 1–30.
[S28] A. Dréau, M. Lesik, L. Rondin, P. Spinicelli, O. Arcizet, J. F. Roch, and V. Jacques, "Avoiding power broadening in optically detected magnetic resonance of single NV defects for enhanced dc magnetic field sensitivity," Physical Review B - Condensed Matter and Materials Physics **84**, 1–8 (2011).
[S29] G de Lange, Z H Wang, D. Riste, V V Dobrovitski, and R Hanson, "Universal Dynamical Decoupling of a Single Solid-State Spin from a Spin Bath," Science **330**, 60–63 (2010).
[S30] Viktor Stepanov and Susumu Takahashi, "Determination of nitrogen spin concentration in diamond using double electron-electron resonance," Physical Review B **94**, 024421 (2016).
[S31] Thomas Wolf, Philipp Neumann, Kazuo Nakamura, Hitoshi Sumiya, Takeshi Ohshima, Junichi Isoya, and Jörg Wrachtrup, "Subpicotesla Diamond Magnetometry," Physical Review X **5**, 041001 (2015).
[S32] Alexander M. Zaitsev, *Optical Properties of Diamond* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2001) p. 502.
[S33] K. Beha, A. Batalov, N. B. Manson, R. Bratschitsch, and A. Leitenstorfer, "Optimum photoluminescence excitation and recharging cycle of single nitrogen-vacancy centers in ultrapure diamond," Physical Review Letters **109**, 1–5 (2012).
[S34] Claire Glover, M. E. Newton, P. M. Martineau, Samantha Quinn, and D. J. Twitchen, "Hydrogen Incorporation in Diamond: The Vacancy-Hydrogen Complex," Physical Review Letters **92**, 135502 (2004).
[S35] Gijs de Lange, Toeno van der Sar, Machiel Blok, Zhi-Hui Wang, Viatcheslav Dobrovitski, and Ronald Hanson, "Controlling the quantum dynamics of a mesoscopic spin bath in diamond," Scientific Reports **2**, 382 (2012).