

Subpicotesla Diamond Magnetometry

Thomas Wolf,^{1,*} Philipp Neumann,^{1,†} Kazuo Nakamura,² Hitoshi Sumiya,³
Takeshi Ohshima,⁴ Junichi Isoya,⁵ and Jörg Wrachtrup¹

¹*3rd Institute of Physics and Research Center SCoPE, University of Stuttgart, 70550 Stuttgart, Germany*

²*Energy System Research Institute, Tokyo Gas Company, Limited, Yokohama 230-0045 Japan*

³*Sumitomo Electric Industries, Limited, Itami, Hyogo 664-0016, Japan*

⁴*Japan Atomic Energy Agency, Takasaki, Gunma 370-1292, Japan*

⁵*Research Center for Knowledge Communities, University of Tsukuba,
1-2 Kasuga, Tsukuba, Ibaraki 305-8550, Japan*

(Received 16 February 2015; revised manuscript received 4 August 2015; published 5 October 2015)

Nitrogen-vacancy (NV) defect centers in diamond are promising solid-state magnetometers. Single centers allow for high-spatial-resolution field imaging but are limited in their magnetic field sensitivity. Using defect-center ensembles, sensitivity can be scaled with \sqrt{N} when N is the number of defects. In the present work, we use an ensemble of $N \sim 10^{11}$ defect centers within an effective sensor volume of $8.5 \times 10^{-4} \text{ mm}^3$ for sensing at room temperature. By carefully eliminating noise sources and using high-quality diamonds with large NV concentrations, we demonstrate, for such sensors, a sensitivity scaling as $1/\sqrt{t}$, where t is the total measurement time. The associated photon-shot-noise-limited magnetic-field sensitivity for ac signals of $f = 20 \text{ kHz}$ is $0.9 \text{ pT}/\sqrt{\text{Hz}}$. For a total measurement time of 100 s, we reach a standard deviation of about 100 fT. Further improvements using decoupling sequences and material optimization could lead to $\text{fT}/\sqrt{\text{Hz}}$ sensitivity.

DOI: [10.1103/PhysRevX.5.041001](https://doi.org/10.1103/PhysRevX.5.041001)

Subject Areas: Optoelectronics, Quantum Physics,
Spintronics

I. EXPERIMENT

Magnetic sensors find application in various areas of science, technology, and medicine [1,2]. Next to robust and widely applicable Hall sensors [3], other methods have been developed that enable new fields of applications and provide insight into aspects of fundamental physics. There are a variety of different sensor implementations, differing in sensitivity, magnetic-field range, measurement bandwidth, operating temperature, and power consumption. For example, sensors based on magnetoresistive effects are well known for their application in hard disk drives. Recent approaches exploit the influence of magnetostrictive materials on high- Q optical cavities [4]. Furthermore, superconducting quantum-interference devices (SQUID) [5] and atomic vapor cells [6] are at the forefront of magnetic sensitivities, and they explore the limits for detection of fundamental flux quantization.

For *detection and localization* of weak magnetic-field sources, sensor size is a further critical parameter as dipolar

magnetic fields decay with $1/r^3$. Consequently, a number of approaches strive for high sensitivity in combination with reduced sensor sizes. For example, SQUIDs have reached dimensions of around 100 nm for scanning probe applications under cryogenic and vacuum conditions [7]. Atomic-vapor magnetometers have been scaled down to 1 mm^3 with sensitivity of $5 \text{ fT}/\sqrt{\text{Hz}}$ for room-temperature applications [8]. While further miniaturization of the latter room-temperature sensors is a current research field, surprisingly, at the nanometer scale, sensors do exist that are capable of detecting spins down to the single level in a reasonable amount of time [9–12]. However, scaling of these sensors towards micron scale and accompanied improved sensitivity has proven to be challenging, especially for spin-based sensors like diamond defects. Here, we demonstrate a magnetic-field sensor using an ensemble of about 10^{11} nitrogen-vacancy (NV) centers in high-quality diamond with an active sensor volume of 850 pl, for which the size-dependent scaling of sensitivity was retained. To this end, we have applied ac magnetometry based on spin-echo techniques, which suppresses low-frequency noise (e.g., temperature, pressure, magnetic-field drifts, and inhomogeneities across the sensor spin ensemble). In addition, we have identified laser and microwave fluctuations as main noise sources and designed appropriate measurement sequences, which act as an effective noise rejection.

* t.wolf@physik.uni-stuttgart.de

† p.neumann@physik.uni-stuttgart.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Diamond magnetometry has been reviewed recently [13]. In short, the favorable material properties of diamond, as well as the optical and spin properties of NV defect centers, allow for optical polarization, manipulation, and readout of its spin state [13] under a wide range of conditions, including temperatures from cryogenic to above 600 K [14], pressures from 0 to GPa [15], and magnetic fields up to about 10 mT [16] for vector magnetometry and >4 T for oriented fields [17–19]. This opens new ways for the implementation of robust solid-state sensors for a variety of quantities [20,21] and fields of applications. In particular, as magnetic-field sensors, NV-based approaches offer the opportunity for detection of magnetic-field signals both with high spatial accuracy (nanometer) and high field sensitivity [12,22]. In this work, we focus on the sensitivity and sensitivity scaling of magnetic-field measurements with ensembles of NV centers [23]. Magnetic-field detection is based on ground-state Zeeman shifts of spin sublevels of NV centers $\Delta E = \gamma \hbar B$, where $\gamma = 2\pi \times 28$ GHz/T is

the gyromagnetic ratio of the electron spin and B is the field to be measured. ΔE is best determined by exploiting coherent control of the electronic spin state of the NV centers in its ground state [Figs. 1(a) and 1(b)]. In essence, the spin acquires a phase $\varphi = \gamma \times B \times T_\varphi$ during sensing time T_φ (B is the averaged field) in Ramsey-type or spin-echo-type measurements [24]. Optical excitation with a laser pulse concludes a single-field evaluation step by invoking spin-state-dependent fluorescence and reinitializing the spin state via the spin selective singlet decay of the NV centers [13]. The fluorescence response is thereby modulated with $\sin \varphi(B)$.

In general, the sensitivity of a magnetic-field measurement is given by $B(t) = \sigma(t)/(dS/dB)$, where the standard deviation of the sensors signal $\sigma(t)$ is compared to the response of the system dS in a changing magnetic field dB . For the particular case of NV centers using a pulsed detection scheme with discrete readout steps, the sensitivity is written as

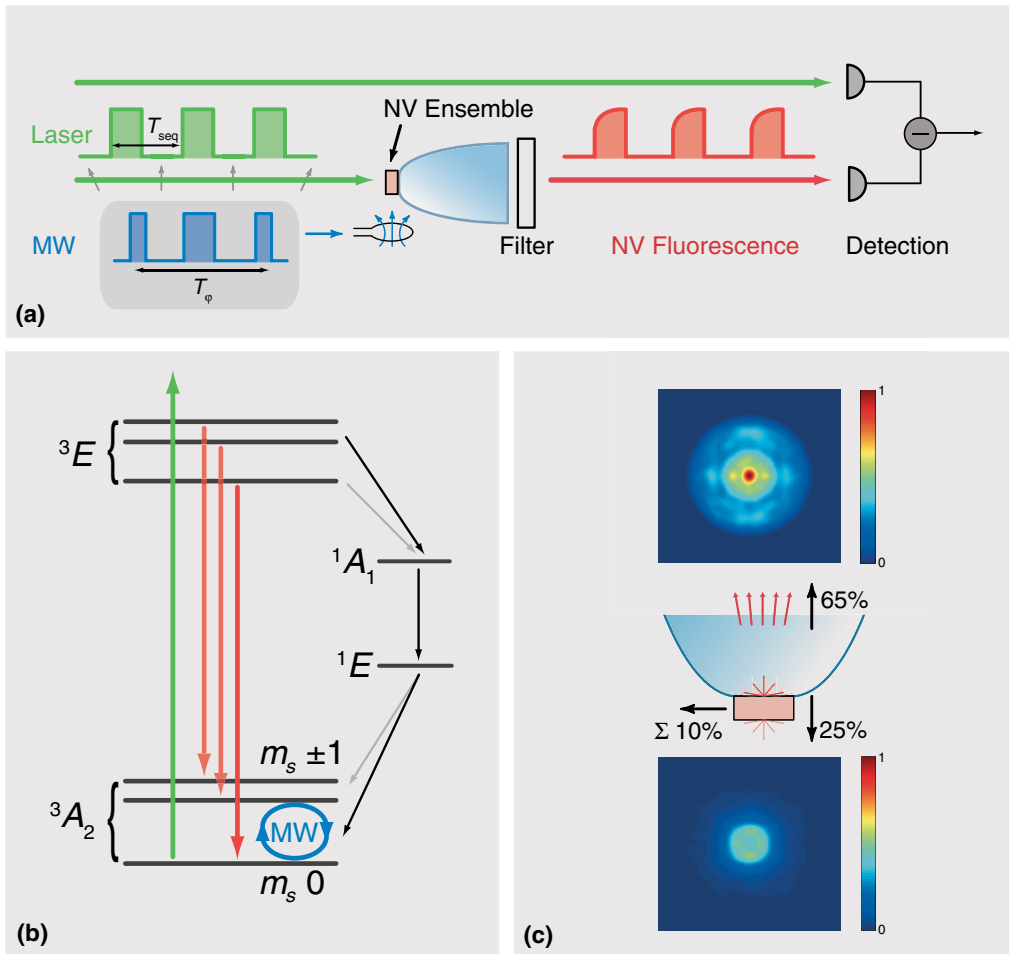


FIG. 1. NV diamond ensemble magnetometer. (a) Pulsed NV experiment: NV ensemble is excited by 532-nm laser pulses. Long-pass-filtered fluorescence is collected with part of the exciting light on a balanced detector. Microwave (MW) pulses are used for NV-spin manipulation. (b) NV-energy-level scheme: Manipulation of electron spin in a triplet ground state. Spin-state-dependent fluorescence allows readout of the spin state. (c) Fluorescence collection with parabolic collector (simulation results).

$$B_{\min}(t) = \frac{\sigma_1}{\gamma \cdot A \cdot T_\phi \sqrt{n}}. \quad (1)$$

Here, $n = t/T_{\text{seq}}$ is the number of field evaluations for a total measurement time t with T_{seq} , σ_1 , and T_ϕ the duration, the standard deviation, and the phase accumulation time of an individual field evaluation, respectively. Parameter A is the system-specific amplitude of the signal modulation [24].

Before dwelling on the accuracy of ensemble NV magnetometry, it is instructive to analyze single-spin measurements. The standard deviation of single-spin sensor readouts σ_1 is dominated by shot noise of the fluorescence signal. Its ultimate limit, however, is spin projection noise due to the statistical nature of the quantum-mechanical readout of the spin state. It is only reached by reducing the relative fluorescence shot noise below the spin-projection noise limit. Steps towards this goal are, for instance, improved fluorescence detection efficiency by waveguiding effects as shown in Ref. [25], repetitive readout [26,27], or different detection schemes [28,29]. Since both fluorescence signal and spin projection are sources of uncorrelated noise, sensitivity scales as \sqrt{n} over a wide range of measurement times, where n is the number of individual sensor readouts.

We now turn to ensemble magnetometry. To further improve sensitivity, σ_1 is decreased with increasing fluorescence signal intensity when measuring on ensembles of NV centers. For independent emitters, σ_1 should scale as $1/\sqrt{N}$, where N is the number of defects contributing to field measurement. Eventually, we calculate the spin-projection-limited magnetic-field sensitivity as

$$B_{\text{QPN}}(t) = \frac{1}{\gamma \sqrt{Nt/T_{\text{seq}}} \cdot T_\phi \cdot e^{-\delta(T_\phi)}} \quad (2)$$

with $e^{-\delta(T_\phi)}$ describing the decay of spin coherence and N the number of NV centers contributing to the signal. The equation is equivalent to the general derivation of Chin *et al.* [30] for the case $T_{\text{seq}} \rightarrow T_\phi$. In this case and for an exponential decay of spin coherence with a time constant T_2 , the minimum of Eq. (2) is achieved for $T_\phi = T_2/2$, which simplifies to $B_{\text{QPN}}(t) = \sqrt{2}e/(\gamma \sqrt{NtT_2}) = 1.3 \times 10^{-11} \text{ T}/\sqrt{NtT_2}$.

Our sensor consists of a 0.9-ppm NV⁻ HPHT diamond, starting from 3-ppm nitrogen before conversion to NV by electron irradiation, and has a thickness of 500 μm with $\langle 111 \rangle$ front planes. Fluorescence from the NV centers after pulsed excitation using a green laser is measured on one channel of a balanced detector. The second channel of the detector is illuminated by part of the green excitation beam split from the exciting laser [Fig. 1(a)]. Measurements shown were conducted at room temperature with a constant offset magnetic field of 46 Gauss along one NV orientation

used for the measurements. The sensitivity of single-center magnetometry is limited by the number of detectable photons from a single defect being given by the photo-physics of the defect. In ensemble magnetometry, an additional challenge is to detect as many fluorescence photons as possible from a finite sample volume. As a result, different collection as well as absorption schemes have been proposed [25,26,28,29,31]. In our approach, we maximize fluorescence signal intensity from the NV centers by using a parabolic-shaped glass lens contacting one side of the diamond. As the étendue of a light source is conserved (for constant intensity), this structure essentially trades the initially large solid angle of fluorescence radiation against the size of the emitting surface area of the structure. Simulations of the structure (see Sec. II) show a collection efficiency higher than 60% [Fig. 1(c)].

Our magnetic-field measurement scheme comprises three steps. First, the NV sensor spins are polarized with a laser pulse. After initialization, we apply a microwave preparation sequence for the B -field measurement. The fluorescence signal is triggered and read out subsequently by launching another laser pulse. Microwave pulses are implemented using a coil antenna. Typical Rabi frequencies are on the order of 5 MHz. For excitation, we focus 400 mW of laser power with a diameter of 47 μm onto the sample. The maximum intensity used is about 10 kW/cm², which is below saturation (about 100 kW/cm²). The sample volume, and hence the number N of defects contributing to the fluorescence signal, is determined by the optical excitation and detection volume. Based on a measurement of the excitation area using a CCD camera and given the collection property of our parabolic lens, which is, at first order, spatially nonselective towards the fluorescence created, we calculate an effective detection volume of $8.5 \times 10^{-4} \text{ mm}^3$. With a density of 0.9 ppm, we estimate that 1.4×10^{11} NV centers contribute to the sensor signal. From Eq. (1), we estimate an expected maximum sensitivity of 100 fT/Hz^{1/2} with $T_\phi = 50 \mu\text{s}$ if we scale values of single NV sensitivities with the mentioned number of NV centers. From Eq. (2), we calculate the spin-projection-noise limit to be 6 fT/Hz^{1/2} (T_ϕ is again set to 50 μs). This estimate relies on the assumption that the results of single readout steps show a normal distribution around a constant mean value (central-limit theorem). This condition is usually met for measurements on single NV centers with comparably small numbers of total signal photons dominated by optical shot noise or spin projection noise—a frequency-independent, uncorrelated white-noise background. Ensemble magnetometry, however, dealing with much higher fluorescence intensities, is plagued by other, correlated and time-dependent noise sources. Since preparation and readout of the measurement relies on discrete preparation steps using laser and microwave pulses, it is essential to analyze the influence of each of these sources on the sensitivity and mitigate their impact.

We start by performing an ac-magnetometry experiment, as demonstrated previously [23,32]. To this end, a spin-echo measurement with pulses $(\pi/2)_x(\pi)_x(\pi/2)_y$ is phase-locked to a sinusoidal ac magnetic field [see Fig. 2(a)], which we intend to sense. We use a phase accumulation time of $T = 50 \mu\text{s}$, and the overall single sequence length is $T_{\text{seq}} = 160 \mu\text{s}$. The phase shift of the last $\pi/2$ pulse in the echo sequence assures maximum sensitivity already for the smallest amplitude of the test field. By increasing the ac amplitude, we increase the accumulated phase linearly with a concomitant sinusoidal fluorescence response. For the following investigations on the reproducibility of individual sensor readouts and its scaling behavior with averaging time, we chose the point of maximum field sensitivity.

Figure 2(b) shows the scaling of the Allan deviation (see Sec. II) of the readout signal of two different measurements. The upper curve corresponds to the spin-echo measurement described above. The second stems from

an identical measurement but without applying microwave pulses in between laser readout pulses. While the Allan deviation may not be considered a valid estimator for the scaling of magnetic-field sensitivity, it provides information on the correlation of consecutive measurements (here, $t_{\text{corr}} < 100 \text{ ms}$) without being affected by overall (long-term) drift, as in the case of standard deviation. As apparent from our measurements, once microwave pulses are applied, sensitivity scales worse than \sqrt{n} . As a result, no further improvement by averaging is achieved. Note that the Allan deviation in conjunction with the applied spin echo-sequence on short time scales does not reveal long-term drifts, e.g., due to changes in temperature [20] or magnetic background field. Thus, only external magnetic noise and variations of the implemented microwave sequence remain obvious culprits.

The impact of microwave amplitude and frequency noise on measurement error depends on the particular choice of

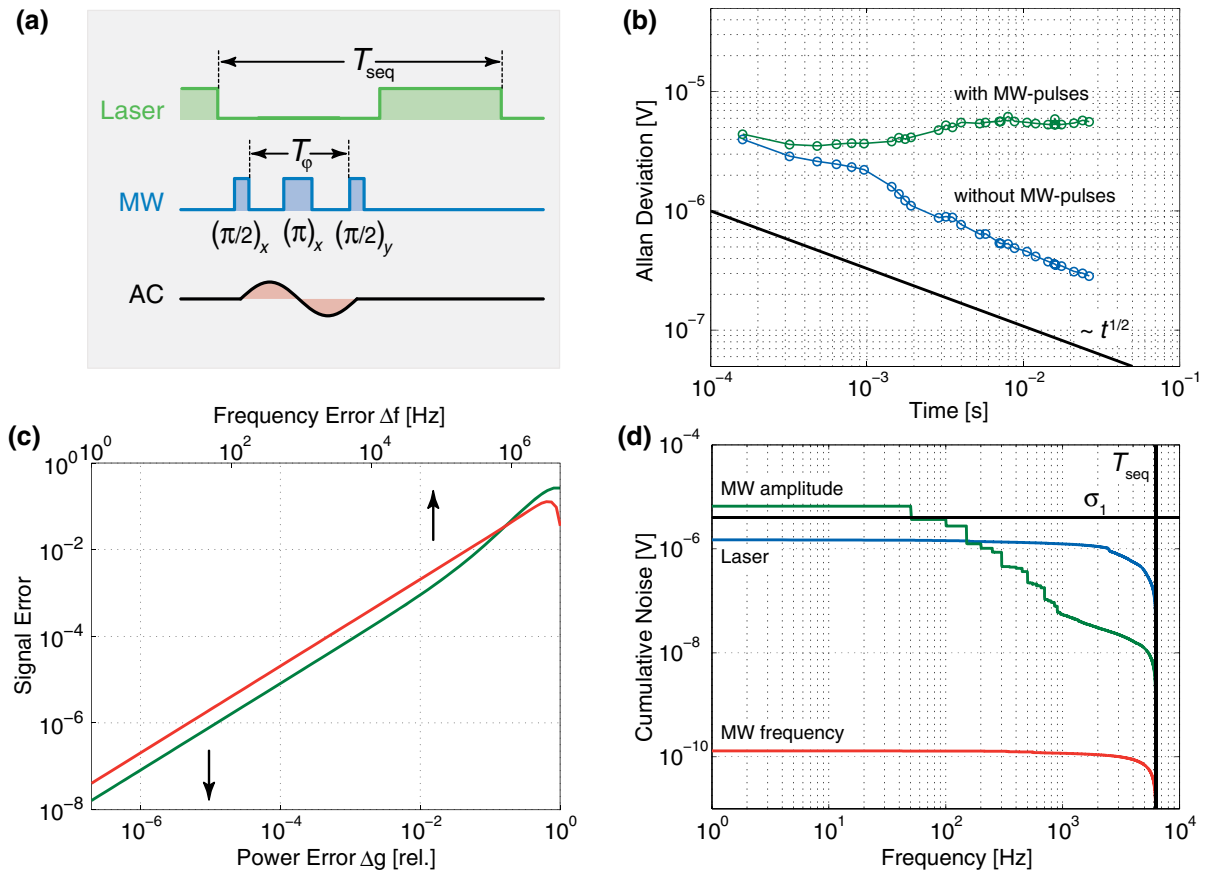


FIG. 2. Influence of non-white noise on magnetometer sensitivity. (a) ac magnetic-field measurement scheme with pulsed sensor readout. (b) Scaling of Allan deviation from the Hahn-echo sequence with (green line) and without (blue line) microwave pulses. The slope of the black line indicates the desired scaling behavior to approach a central limit. If microwave pulses are applied, scaling is worse than \sqrt{t} . (c) Calculated spin-state population error (Δz) after the Hahn-echo sequence over relative microwave power error Δg (green line) and frequency error Δf (red line). (d) Cumulative noise over sequence length T_{seq} from high to low frequencies of laser (blue line), microwave power (green line), and microwave frequency (red line). The measured deviation of single readout steps σ_1 is indicated by the black horizontal line. Varying microwave power is expected to dominate the distribution of magnetic-field evaluations on longer time scales. Therefore, results of field evaluations do not share a common central limit over time.

the microwave-pulse sequence and the noise frequency. For noise correlation times longer than the length of a single readout sequence (T_{seq}), an error in microwave amplitude or frequency can be taken to be constant throughout a single sequence. In this regime, the signal error after a microwave-pulse sequence due to a set of amplitude and frequency error parameters can be extracted from simulations of coherent spin rotations under the NV-spin Hamiltonian [33,34]. Figure 2(c) shows results for signal error scaling for two limiting cases: (1) scaling with the relative microwave power error Δg ($\Delta f = 0$) and (2) scaling with the absolute microwave frequency error Δf ($\Delta g = 0$). Both quantities show a linear scaling behavior in the relevant parameter range.

We measured amplitude and phase noise of the microwave system using established cross-correlation techniques [35]. Both are given for frequencies f below the inverse sequence length ($1/T_{\text{seq}}$) in Fig. 2(d) (cumulative root sum of squares from $1/T_{\text{seq}}$ to f). The optical noise is given in the same way, and the Allan deviation of single-sequence readouts is indicated as the horizontal black line. Three results can be inferred from the graph. First, the influence of microwave-frequency noise on signal scaling is negligible. Second, the optical low-frequency noise ($f < 1/T_{\text{seq}}$) is well below the noise of single-sequence readouts, leading to \sqrt{n} scaling when no microwave pulses were applied. Finally, the effect of microwave amplitude noise is small on the time scale of single readout steps but increases towards longer time scales. In this case, scaling of the measurements is worse than \sqrt{n} , as the central limit theorem does not apply. In essence, varying microwave amplitudes lead to improper conversion of field-to-signal amplitude that prevents \sqrt{n} scaling.

One way to reduce the impact of inaccurate microwave pulses on scaling of the readout signal is to reference the signal on a time scale shorter than the characteristic correlation time of the noise. Owing to the photo-physical dynamics of the NV center, the spin signal is typically read out in the first part of a laser pulse. It is common practice to reference this signal [“1” in Fig. 3(a)] to the steady-state fluorescence level after reinitialization of the NV centers at the end of the laser pulse [“2” in Fig. 3(a)] yielding the measurement signal S_B . Implicitly, this procedure mitigates optical noise with a correlation time longer than the laser-pulse length (100 μs)—see Sec. II for an analytic derivation of the noise reduction. The described procedure only affects the laser-induced correlated fluctuations of the signal. A reference for the state preparation with microwave pulses can be established likewise. By introducing a second preparation and readout sequence as shown in Fig. 3(a) for signal S_D , microwave-induced noise is reduced. To understand the impact of this procedure, we calculate the noise reduction for different ways of measuring the NV signal (S_A to S_D) as shown in Fig. 3(a). Signal S_A does not contain any referencing. S_B , as explained above,

implements one filtering step for the optical signal on the time scale of the laser-pulse length. S_C results in one referencing step for both optical and microwave-related noise on the time scale of the sequence length. Finally, S_D gives two referencing steps versus the optical part and one for the microwave contribution. Calculation of noise reduction for signals $S_{A/C/D}$ is analogous to signal S_B . The calculated noise-reduction functions are given in Fig. 3(b) for the same set of parameters used for the measurements from above, $t_L = 100 \mu\text{s}$, $\Delta t = 10 \mu\text{s}$, and $T_{\text{seq}} = 160 \mu\text{s}$. The calculation is restricted to noise frequencies in the low-frequency regime up to the range of $1/T_{\text{seq}}$. Note that the response of the measured signal corresponding to the implementation of the signal measurement procedure S_D results in filter X_D for optical noise because of two referencing steps. It results in filter X_C for the microwave-noise contributions since state preparation by microwave only affects the signal in the first part of the laser pulses.

With the calculated noise, we again give the cumulative noise (root sum of squares) weighted with the respective filters for microwave components and the optical part corresponding to the measurement procedure S_D from two consecutive readouts [Fig. 3(c)]. The calculation shows that referencing the signal in this way very efficiently suppresses low-frequency noise components such that \sqrt{t} scaling is achieved.

Next, we repeat the ac-magnetometry sequence as described above and shown in Fig. 2(a); however, we measure signals S_B and S_D . For this procedure, we apply a magnetic test field and retrieve the amplitude of the sensor response [Fig. 3(d)]. We then switch off the actual ac field to (1) reach the working point of highest sensitivity and (2) exclude additional noise sources from the ac signal itself. The scaling of the sensitivity, given here as the standard deviation for the two signals S_B and S_D , is shown in Fig. 3(e) as a function of the number of repetitions (or measurement time t). In the case of signal S_B (blue line), increased averaging time does not improve the measurement result. Signal S_D (green line) shows a scaling with $1/\sqrt{t}$, where t is the total signal averaging time, which reaches a sensitivity of 0.9 pT/Hz^{1/2}. For longer measurement time, an absolute sensitivity of around 100 fT is achieved. Our measurements are not limited by temperature variation nor by external in-band magnetic noise. Experimentally, we find that the standard deviation of the signal S_D is, by a factor of 6, above the fluorescence shot-noise level, which we would expect for a simple readout of the repolarization signal only (S_A). This result agrees well with the expected increase of uncorrelated noise in our measurement scheme. While correlated noise is largely suppressed with the procedures described, uncorrelated noise increases by a factor $\sqrt{2}$ with every referencing step implemented (if we assume identical noise density for the uncorrelated noise of the two signals

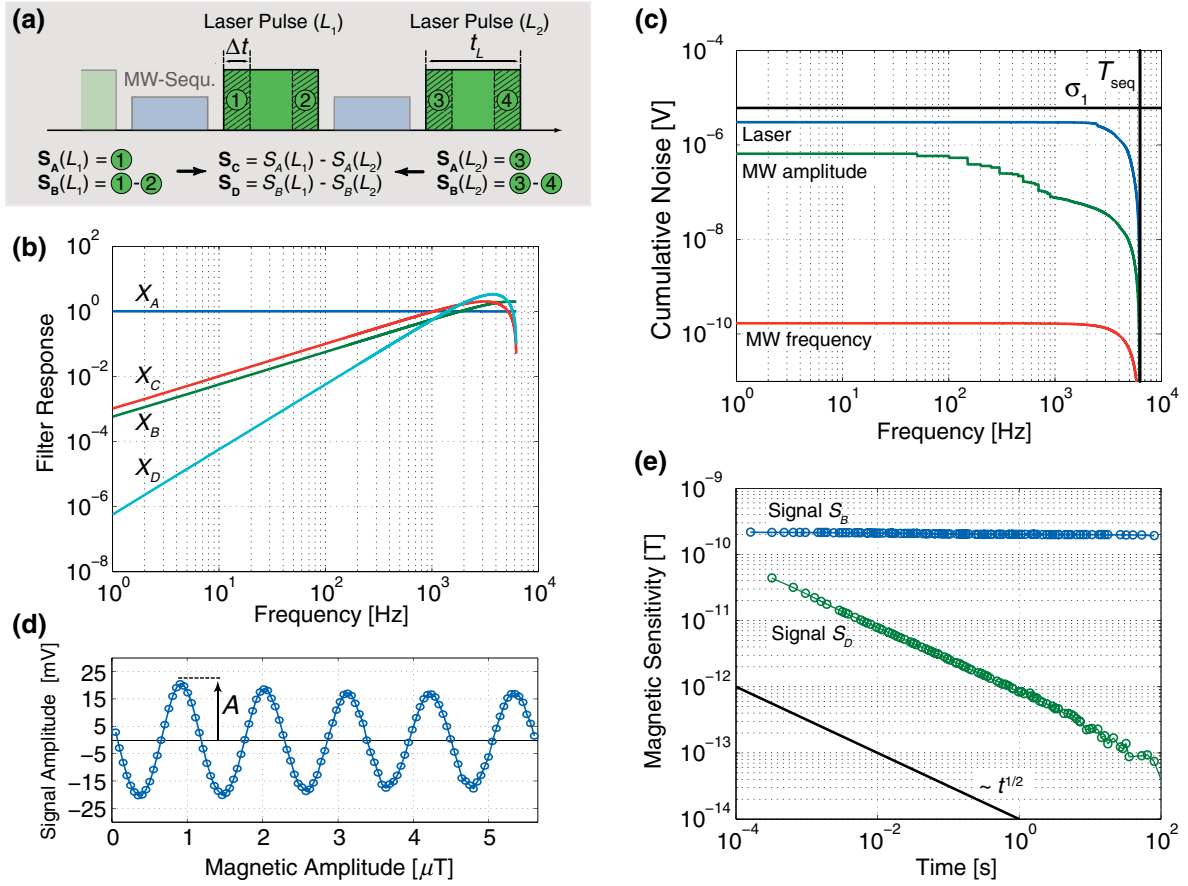


FIG. 3. Sequence filter and sensitivity scaling. (a) Measurement signals ($S_A - S_D$) represent different ways of taking the signal from the sensor. (b) Corresponding sequence-filter functions resulting from signals $S_A - S_D$. (c) Cumulative, sequence-filtered noise of laser (blue line), microwave power (green line), and microwave frequency (red line) for signal S_D . The black horizontal line indicates single-sequence readout deviation (σ_1). σ_1 remains the dominant contribution within the time scale shown. (d) Magnetic measurement of the test field with varying field amplitude retrieving sensor response A [Eq. (1)]. (e) Scaling of magnetic sensitivity (standard deviation over time) of signal S_B (blue line) and S_D (green line). The slope of the black line again indicates the aspired scaling behavior with \sqrt{t} .

referenced). Since we introduced three referencing steps (exciting laser against fluorescence and two in measurement procedure S_D) and an additional factor for doubling the measurement time, we thereby effectively increase the contribution of uncorrelated noise by $\sqrt{2^4} = 4$.

When we extrapolate sensitivity from measurements on single NV centers (\sqrt{N}), we expect a value of $100 \text{ fT}/\sqrt{\text{Hz}}$ [resulting from Eq. (1)] for our sensor. This value deviates by 1 order of magnitude from our experimental findings. We can resolve this discrepancy by accounting for the reduction in contrast when measuring on one of four NV axes and the increase in noise by 6, as mentioned before. We predict that the sensitivity of measurements on NV ensembles can even exceed the projection derived from single NV measurements due to improved fluorescence collection efficiency. Finally, we want to emphasize that the implementation of the measurement procedure S_D , except for a decrease in measurement rate, does not impose any additional restrictions. In particular, we remark that, as we require control of the ac magnetic field (e.g., switch off in

every second measurement, in-phase with spin-echo sequence), also in real measurements, the source of the ac magnetic field needs to be controlled (e.g., invoked flips of electron or nuclear spins to be measured [36,37]).

The present work advances NV ensemble magnetometers to sub-pT/ $\sqrt{\text{Hz}}$ sensitivity, which is particularly remarkable when it comes to sensitivity weighted by sensor size. To this end, we highlight the role of technical noise and its mitigation to achieve *sensitivity scaling* ($\propto \sqrt{t}$, \sqrt{N}) in NV ensemble magnetometry. Different strategies are conceivable to further improve the magnetic sensitivity. In Ref. [38], higher-order dynamical decoupling sequences were applied to NV ensembles yielding a phase memory time of 2 ms, which is the limit set by longitudinal spin relaxation of NV centers at room temperature. With the experimental settings described here, with a similar amount of NV centers and identical efficiency of the filters applied, this would yield a sensitivity of $40 \text{ fT}/\sqrt{\text{Hz}}$. This sensitivity is still almost 2 orders of magnitude above the limit set by the spin projection noise ($0.9 \text{ fT}/\sqrt{\text{Hz}}$). The latter

value itself allows for detection of proton spins in water in a microscopically resolvable volume (10^{-18} m³) in less than one second, bringing new fields of application into reach: for example, minimal-invasive detection of biomagnetic fields and microfluidic chemical analysis, as well as cellular and neuronal imaging, where characteristic dimensions are on the micron scale. Nuclear-spin-assisted repetitive readout [26], infrared-absorption-based readout [28], or enhancement by optical cavities [29] are strategies to reach the projection noise limit.

II. METHODS

A. Experimental details

An existing problem for the implementation of magnetic sensors based on the detection of fluorescence signals from NV ensembles is the efficient collection of fluorescence light created in the diamond. As the refractive index of diamond is about 2.4 in the visible wavelength range, the critical angle for total internal reflection at a diamond-air interface is low ($\alpha_c \approx 24.5^\circ$). In other words, the probability for light to undergo multiple total internal reflections from all surfaces of the diamond is high. In a perfect diamond, there is a high probability that light is trapped without ever leaving the diamond. However, in a realistic situation, scattering at imperfect surfaces and edges, as well as slow changes of reflection angles by slightly nonparallel (“tapered”) opposing surfaces, breaks the symmetry, finally leading to angles of incidence below the critical angle. Contacting a single side of the diamond with a higher refractive index material, e.g., glass, increases α_c on this side. This results in higher output coupling towards this direction and, second, acts as a drain for light reflected multiple times. As the light coupled in the glass is diffracted into a 2π solid angle, we use an optical structure that collects light from close to 2π , independent from the position within the contacted diamond interface. The structure we chose is known in the field of nonimaging optics as a compound parabolic concentrator (CPC) [39]. Essentially, it consists of two parallel, opposing circular surfaces of different sizes that are connected by a segment of a parabolic functional profile rotated around the central (symmetry) axis such that opposing parabola profiles mutually share their foci. By design, the parabolic profile ends towards one side in height at the position of the foci, thereby defining the smaller circular end face where the diamond was contacted using optical adhesive. A suitable parabolic lens was bought from Edmund Optics (#65-441). The NV ensemble sensor was produced from a HPHT diamond containing 3-ppm initial P1 concentration. The diamond was irradiated using 2-MeV electrons to a total dose of 1.8×10^{18} cm², followed by annealing for 5 h at 800 °C. Collection and transfer efficiency of the combined system (diamond+CPC) was simulated using ray-tracing software. Simulation of the diamond geometry

($1.6 \times 1.6 \times 0.5$ -mm side faces tilted by 5° to lift mathematical perfection) results in $\approx 65\%$ of the fluorescence light leaving the parabolic lens towards the detector. We did not attempt to geometrically optimize the collection efficiency in experiment, as the resulting improvement would enter into sensitivity only by a square-root reduced factor.

In experiment, we use a 640-nm long-pass filter followed by two aspheric lenses after the parabolic lens to transfer the light to the detector. We use a home-built, balanced detector with two 1-cm² photodiodes, a responsivity of about 500 V/W, a bandwidth of about 6.5 MHz, and a broadband noise equivalent power approximately equal to 15 pW/ $\sqrt{\text{Hz}}$. In the case of continuous excitation of the diamond with 400-mW laser power at 532 nm (Lighthouse Photonics, Sprout-G), we measure about $P_{\text{fl}} = 28$ mW incident on the detector. Detector signal voltage is measured with a 16-bit digitizer (Gage Applied, CS1622). An acousto-optic modulator (Crystal Technology #3250-220) in combination with a pulse generator (Tektronix DTG5274) was used for laser-pulse generation. We chose to excite the NV ensemble using 100- μs laser pulses. This is rather long compared to the 3- μs pulses used in experiments on single NV centers. Although the NV defects in the center of the Gaussian laser beam experience almost 1/10 of the saturation intensity, defects at the rim of the Gaussian beam do not. There, the intensity is about 1/100 of what is necessary for saturation. In addition, the signal contribution from these NV defects is larger because of their larger number. In this weak excitation regime, the optimum integration period is longer than for single NV centers under saturation conditions (300 ns). Using a longer laser pulse, we wanted to make sure that all NV centers are reinitialized to avoid correlations between consecutive measurement steps. In particular applications, the given experimental conditions need to be considered in order to find the corresponding optimum excitation time.

To relate the shot noise level of the measured (continuous) signal power to the standard deviation of our measurement, the shot noise for the reduced fluorescence collection window in the particular case of our pulsed measurement with $\Delta t/T_{\text{seq}} = 10 \mu\text{s}/160 \mu\text{s}$ is calculated. With the electron charge $e = 1.602 \times 10^{-19}$ C, $I_{\text{fl}} \approx P_{\text{fl}} \cdot 0.5$ A/W the current behind the photodiode, and the detector gain $G = 1000$ V/A, the shot-noise voltage per $\sqrt{\text{Hz}}$ bandwidth reads

$$N = (2eI_{\text{fl}} \cdot \Delta t/T_{\text{seq}})^{1/2} \cdot G \approx 1.7 \times 10^{-8} \text{ V}. \quad (3)$$

The experiment, including the described referencing steps (noise is expected to increase by 4 as described in the text), yields a signal standard deviation of $N = \sigma_1/\sqrt{n} \approx 1 \times 10^{-7}$ V [compare Eq. (1)]. Thus, in experiment, an increase by a factor of 6 instead of the calculated factor 4 was observed.

B. Magnetic-field sensitivity

Having calculated the noise of our sensor, we can now deduce the magnetic-field sensitivity according to Eq. (1). In our case, the derivative of the signal S with respect to the amplitude B_0 of the ac signal yields $dS/dB_0 = A \cdot \gamma \cdot T_\varphi \cdot 2/\pi = 0.02 \text{ V} \cdot 2\pi 28 \text{ GHz/T} \cdot 50 \text{ } \mu\text{s} \cdot 2/\pi = 112000 \text{ V/T}$. Here, A is the measured voltage amplitude of the sinusoidal oscillation shown in Fig. 3(d). Hence, the sensitivity equates to $0.9 \text{ pT}/\sqrt{\text{Hz}}$. Note that the additional factor $\pi/2$ for the sensitivity is due to the derivative with respect to the amplitude B_0 of an ac magnetic field $B_0 \cos(2\pi t/T_\varphi)$, in contrast to the average field strength $2B_0/\pi$. The error in the magnetic-field response A is estimated by the sum of the optical noise when measuring two points within the phase cycle, each read out one time: $\delta A \leq 2 \times 6 \times 10^{-6} \text{ V}$, yielding a relative error $\delta A/A \leq 0.06\%$. We estimate the signal contrast to be $\approx 0.5\%$, which is below the maximum expected value of $30\%/4 = 7.5\%$ when only one NV orientation out of four contributes to the signal contrast.

C. Allan deviation

The (nonoverlapping) Allan deviation $\sigma_A(\tau)$ of a set of data samples $S = [S_1, S_2, \dots, S_n]$ with sampling frequency $1/t'$ is defined for a given time interval τ by

$$\sigma_A^2(\tau) = \frac{1}{2} \langle (x_{i+1} - x_i)^2 \rangle_\tau. \quad (4)$$

In the equation, x_i denotes the mean $\langle \cdot \rangle$ over the subset of $m = \tau/t'$ successive elements of S within the i th τ interval:

$$x_i = \langle [S_{(i-1)m+1}, S_{(i-1)m+2}, \dots, S_{im}] \rangle. \quad (5)$$

D. Error scaling with microwave amplitude and frequency

To estimate the impact of microwave-pulse errors on the measurement, we calculate the population difference between the target state (ideal pulses) and the outcome of a pulse sequence with constant error in microwave

frequency and microwave power throughout a sequence (Hahn echo). Successive coherent spin rotations are calculated using the NV-spin Hamiltonian:

$$H = DS_z^2 + B_z(\gamma S_z + \gamma_N I_z) + AS_z I_z. \quad (6)$$

$D = 2.87 \text{ GHz}$ is the zero-field splitting, $\gamma/2\pi = 28.7 \text{ GHz/T}$ is the gyromagnetic ratio of the NV-electron spin, $\gamma_N/2\pi = 3.08 \text{ MHz/T}$ is the nuclear gyromagnetic ratio of ^{14}N , and $A = 2.16 \text{ MHz}$ is the hyperfine coupling between NV-electron and ^{14}N -nuclear spin. S_z and I_z are the electron and nuclear-spin projection operators, respectively.

E. Noise reduction

Noise reduction X_B within the NV signal by mutual referencing of discrete readout windows [e.g., for signal S_B of Fig. 3(a)] is calculated by the Fourier transform of the respective signal integration window C_B :

$$X_B(\omega) = \left| \int_{-\infty}^{\infty} dt e^{i\omega t} C_B(t) \right|, \quad (7a)$$

$$C_B(t) = \begin{cases} 1 & t \in [0, \Delta t] \\ -1 & t \in [t_L - \Delta t, t_L] \\ 0 & \text{otherwise.} \end{cases} \quad (7b)$$

In the latter, the length of the laser pulse is denoted by t_L and the integration time is given by Δt [see also Fig. 3(a)]. Note that this filter is different from (1) filter functions implemented with microwave pulses with the intention to shape the response of NV centers towards a certain band of magnetic signal frequencies [40,41] and (2) filter functions related to data (post-)processing (e.g., to improve the signal-to-noise ratio). The filter referred to here is an intrinsic part of the measurement system based on the fact that we measure discrete, noncontinuous time frames. Explicit evaluation of Eqs. (7) for the case of signal S_B yields the noise-reduction function:

$$X_B = \sqrt{|2/\omega^2 \cdot [2 - 2 \cos \omega \Delta t + \cos \omega(t_L - 2\Delta t) + \cos \omega t_L - 2 \cos \omega(t_L - \Delta t)]|}. \quad (8)$$

Filter functions $X_{A/C/D}$ are calculated in the same way using the corresponding signal windows $C_{A/C/D}$.

A remark on the choice of microwave sequence: As described above, we chose, in the present work, to reference two equal Hahn-echo sequences in order to suppress noise related to the NV state preparation with microwave pulses. As an alternative, two Hahn-echo sequences with the rotation axes of their final $\pi/2$

pulses differing by 180° $[(\pi/2)_x - (\pi)_x - (\pi/2)_y$ and $(\pi/2)_x - (\pi)_x - (-\pi/2)_y]$ may as well lead to efficient noise suppression without the need to turn off the signal in every second sequence. In this work, however, we decided to use two completely identical pulse sequences (except for the ac signal part) such that any error or slow fluctuation on the time scale of two consecutive echos is rejected. In the protocol suggested above, sources of errors such as slowly

varying phase differences between $-y$ and y pulses will still lead to increased noise of the optical signal. In the case of a particular application, one again has to check the present noise sources and consider proper counter measures. It might very well be the case that, under some circumstances, relaxed requirements to the degree of noise reduction are sufficient.

ACKNOWLEDGMENTS

We thank Bernhard Grotz, Marco Di Sarno, and Steffen Steinert for fruitful discussions. We acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) via Project Nos. SFB/TRR 21 and No. FOR 1493, and together with the Japan Science and Technology Agency (JST) for funding via their joint research programme No. FOR 1482. Furthermore, we received funding from the European Union (DIADEMS, SIQS, ERC SQUTEC) and from the Max-Planck Society.

-
- [1] R. Hari and R. Salmelin, *Magnetoencephalography: From SQUIDS to Neuroscience: Neuroimage 20th Anniversary Special Edition*, *NeuroImage* **61**, 386 (2012).
- [2] C. Gaffney, *Detecting Trends in the Prediction of the Buried Past: A Review of Geophysical Techniques in Archaeology*, *Archaeometry* **50**, 313 (2008).
- [3] P. Ripka and M. Janosek, *Advances in Magnetic Field Sensors*, *IEEE Sens. J.* **10**, 1108 (2010).
- [4] S. Forstner, E. Sheridan, J. Knittel, C. L. Humphreys, G. A. Brawley, H. Rubinsztein-Dunlop, and W. P. Bowen, *Ultrasensitive Optomechanical Magnetometry*, *Adv. Mater.* **26**, 6348 (2014).
- [5] D. Drung, C. Assmann, J. Beyer, A. Kirste, M. Peters, F. Ruede, and T. Schurig, *Highly Sensitive and Easy-to-Use SQUID Sensors*, *IEEE Trans. Appl. Supercond.* **17**, 699 (2007).
- [6] H. B. Dang, A. C. Maloof, and M. V. Romalis, *Ultrahigh Sensitivity Magnetic Field and Magnetization Measurements with an Atomic Magnetometer*, *Appl. Phys. Lett.* **97**, 151110 (2010).
- [7] A. Finkler, D. Vasyukov, Y. Segev, L. Ne'eman, E. O. Lachman, M. L. Rappaport, Y. Myasoedov, E. Zeldov, and M. E. Huber, *Scanning Superconducting Quantum Interference Device on a Tip for Magnetic Imaging of Nanoscale Phenomena*, *Rev. Sci. Instrum.* **83**, 073702 (2012).
- [8] W. C. Griffith, S. Knappe, and J. Kitching, *Femtotesla Atomic Magnetometry in a Microfabricated Vapor Cell*, *Opt. Express* **18**, 27167 (2010).
- [9] D. Rugar, C. S. Yannoni, and J. A. Sidles, *Mechanical Detection of Magnetic Resonance*, *Nature (London)* **360**, 563 (1992).
- [10] C. L. Degen, M. Poggio, H. J. Mamin, C. T. Rettner, and D. Rugar, *Nanoscale Magnetic Resonance Imaging*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1313 (2009).
- [11] M. E. Huber, N. C. Koshnick, H. Bluhm, L. J. Archuleta, T. Azua, P. G. Björnsson, B. W. Gardner, S. T. Halloran, E. A. Lucero, and K. A. Moler, *Gradiometric Micro-SQUID Susceptometer for Scanning Measurements of Mesoscopic Samples*, *Rev. Sci. Instrum.* **79**, 053704 (2008).
- [12] G. Balasubramanian, I. Y. Chan, R. Kolesov, M. Al-Hmoud, J. Tisler, C. Shin, C. Kim, A. Wojcik, P. R. Hemmer, A. Krueger, T. Hanke, A. Leitenstorfer, R. Bratschitsch, F. Jelezko, and J. Wrachtrup, *Nanoscale Imaging Magnetometry with Diamond Spins under Ambient Conditions*, *Nature (London)* **455**, 648 (2008).
- [13] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. L. Hollenberg, *The Nitrogen-Vacancy Colour Centre in Diamond*, *Phys. Rep.* **528**, 1 (2013).
- [14] D. M. Toyli, D. J. Christle, A. Alkauskas, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom, *Measurement and Control of Single Nitrogen-Vacancy Center Spins above 600 K*, *Phys. Rev. X* **2**, 031001 (2012).
- [15] M. W. Doherty, V. V. Struzhkin, D. A. Simpson, L. P. McGuinness, Y. Meng, A. Stacey, T. J. Karle, R. J. Hemley, N. B. Manson, L. C. L. Hollenberg, and S. Prawer, *Electronic Properties and Metrology Applications of the Diamond NV(-) Center under Pressure*, *Phys. Rev. Lett.* **112**, 047601 (2014).
- [16] S. Steinert, F. Dolde, P. Neumann, A. Aird, B. Naydenov, G. Balasubramanian, F. Jelezko, and J. Wrachtrup, *High Sensitivity Magnetic Imaging Using an Array of Spins in Diamond*, *Rev. Sci. Instrum.* **81**, 043705 (2010).
- [17] G. Waldherr, J. Beck, P. Neumann, R. S. Said, M. Nitsche, M. L. Markham, D. J. Twitchen, J. Twamley, F. Jelezko, and J. Wrachtrup, *High-Dynamic-Range Magnetometry with a Single Nuclear Spin in Diamond*, *Nat. Nanotechnol.* **7**, 105 (2011).
- [18] V. Stepanov, F. H. Cho, C. Abeywardana, and S. Takahashi, *High-Frequency and High-Field Optically Detected Magnetic Resonance of Nitrogen-Vacancy Centers in Diamond*, *Appl. Phys. Lett.* **106**, 063111 (2015).
- [19] N. Aslam, M. Pfender, R. Stöhr, P. Neumann, M. Scheffler, H. Sumiya, H. Abe, S. Onoda, T. Ohshima, J. Isoya, and J. Wrachtrup, *Single Spin Optically Detected Magnetic Resonance with E-Band Microwave Resonators*, *arXiv:1503.04134*.
- [20] 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*, *Phys. Rev. Lett.* **104**, 070801 (2010).
- [21] M. W. Doherty, J. Michl, F. Dolde, I. Jakobi, P. Neumann, N. B. Manson, and J. Wrachtrup, *Measuring the Defect Structure Orientation of a Single NV Centre in Diamond*, *New J. Phys.* **16**, 063067 (2014).
- [22] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. Gurudev Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin, *Nanoscale Magnetic Sensing with an Individual Electronic Spin in Diamond*, *Nature (London)* **455**, 644 (2008).
- [23] 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*, *Phys. Rev. B* **80**, 115202 (2009).

- [24] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin, *High-Sensitivity Diamond Magnetometer with Nanoscale Resolution*, *Nat. Phys.* **4**, 810 (2008).
- [25] D. Le Sage, L. M. Pham, N. Bar-Gill, C. Belthangady, M. D. Lukin, A. Yacoby, and R. L. Walsworth, *Efficient Photon Detection from Color Centers in a Diamond Optical Waveguide*, *Phys. Rev. B* **85**, 121202 (2012).
- [26] L. Jiang, J. S. Hodges, J. R. Maze, P. Maurer, J. M. Taylor, D. G. Cory, P. R. Hemmer, R. L. Walsworth, A. Yacoby, A. S. Zibrov, and M. D. Lukin, *Repetitive Readout of a Single Electronic Spin via Quantum Logic with Nuclear Spin Ancillae*, *Science* **326**, 267 (2009).
- [27] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko, *Single-Shot Readout of a Single Nuclear Spin*, *Science* **329**, 542 (2010).
- [28] V. M. Acosta, E. Bauch, A. Jarmola, L. J. Zipp, M. P. Ledbetter, and D. Budker, *Broadband Magnetometry by Infrared-Absorption Detection of Nitrogen-Vacancy Ensembles in Diamond*, *Appl. Phys. Lett.* **97**, 174104 (2010).
- [29] K. Jensen, N. Leefer, A. Jarmola, Y. Dumeige, V. M. Acosta, P. Kehayias, B. Patton, and D. Budker, *Cavity-Enhanced Room-Temperature Magnetometry Using Absorption by Nitrogen-Vacancy Centers in Diamond*, *Phys. Rev. Lett.* **112**, 160802 (2014).
- [30] A. W. Chin, S. F. Huelga, and M. B. Plenio, *Quantum Metrology in Non-Markovian Environments*, *Phys. Rev. Lett.* **109**, 233601 (2012).
- [31] H. Clevenson, M. E. Trusheim, T. Schroder, C. Teale, D. Braje, and D. Englund, *Broadband Magnetometry and Temperature Sensing with a Light Trapping Diamond Waveguide*, [arXiv:1406.5235](https://arxiv.org/abs/1406.5235).
- [32] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, *Ultralong Spin Coherence Time in Isotopically Engineered Diamond*, *Nat. Mater.* **8**, 383 (2009).
- [33] H. K. Cummins, G. Llewellyn, and J. A. Jones, *Tackling Systematic Errors in Quantum Logic Gates with Composite Rotations*, *Phys. Rev. A* **67**, 042308 (2003).
- [34] W. G. Breiland, H. C. Brenner, and C. B. Harris, *Coherence in Multilevel Systems. I. Coherence in Excited States and Its Application to Optically Detected Magnetic Resonance in Phosphorescent Triplet States*, *J. Chem. Phys.* **62**, 3458 (1975).
- [35] E. Rubiola and F. Vernotte, *The Cross-Spectrum Experimental Method*, [arXiv:1003.0113](https://arxiv.org/abs/1003.0113).
- [36] P. Neumann, R. Kolesov, B. Naydenov, J. Beck, F. Rempp, M. Steiner, V. Jacques, G. Balasubramanian, M. L. Markham, D. J. Twitchen, S. Pezzagna, J. Meijer, J. Twamley, F. Jelezko, and J. Wrachtrup, *Quantum Register Based on Coupled Electron Spins in a Room-Temperature Solid*, *Nat. Phys.* **6**, 249 (2010).
- [37] H. J. Mamin, M. Kim, M. H. Sherwood, C. T. Rettner, K. Ohno, D. D. Awschalom, and D. Rugar, *Nanoscale Nuclear Magnetic Resonance with a Nitrogen-Vacancy Spin Sensor*, *Science* **339**, 557 (2013).
- [38] L. M. Pham, N. Bar-Gill, C. Belthangady, D. Le Sage, P. Cappellaro, M. D. Lukin, A. Yacoby, and R. L. Walsworth, *Enhanced Solid-State Multispin Metrology Using Dynamical Decoupling*, *Phys. Rev. B* **86**, 045214 (2012).
- [39] W. T. Welford and R. Winston, *The Optics of Nonimaging Concentrators: Light and Solar Energy* (Academic Press, New York, 1978).
- [40] B. Naydenov, F. Dolde, L. T. Hall, C. Shin, H. Fedder, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup, *Dynamical Decoupling of a Single-Electron Spin at Room Temperature*, *Phys. Rev. B* **83**, 081201 (2011).
- [41] G. de Lange, Z. H. Wang, D. Rist, V. V. Dobrovitski, and R. Hanson, *Universal Dynamical Decoupling of a Single Solid-State Spin from a Spin Bath*, *Science* **330**, 60 (2010).