Impact of microwave phase noise on diamond quantum sensing

Andris Berzins,[1,*] Maziar Saleh Ziabari,[1,*] Yaser Silani,[1] Ilja Fescenko,[1,2] Joshua T. Damron,[1,3] John F. Barry,[4] Andrey Jarmola,[5,6] Pauli Kehayias,[4,7] Bryan A. Richards,[1] Janis Smits,[1,†] and Victor M. Acosta[1,‡]

[1]*Center for High Technology Materials and Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87106, USA*
[2]*Laser Centre, Faculty of Science and Technology, University of Latvia, Riga LV-1004, Latvia*
[3]*Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA*
[4]*Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02421, USA*
[5]*ODMR Technologies Inc., El Cerrito, California 94530, USA*
[6]*Department of Physics, University of California, Berkeley, California 94720, USA*
[7]*Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

(Received 9 July 2024; revised 23 July 2024; accepted 9 October 2024; published 15 November 2024)

Precision optical measurements of the electron-spin precession of nitrogen-vacancy (NV) centers in diamond form the basis of numerous applications. The most sensitivity-demanding applications, such as femtotesla magnetometry, require the ability to measure changes in GHz spin transition frequencies at the sub-millihertz level, corresponding to a fractional resolution of better than 10^{-12}. Here we study the impact of microwave (MW) phase noise on the response of an NV sensor. Fluctuations of the phase of the MW waveform cause undesired rotations of the NV spin state. These fluctuations are imprinted in the optical readout signal and, left unmitigated, are indistinguishable from magnetic-field noise. We show that the phase noise of several common commercial MW generators results in an effective $\mathrm{pT\,s^{1/2}}$-range noise floor that varies with the MW carrier frequency and the detection frequency of the pulse sequence. The data are described by a frequency-domain model incorporating the MW phase-noise spectrum and the filter-function response of the sensing protocol. For controlled injection of white and random-walk phase noise, the observed NV magnetic noise floor is described by simple analytic expressions that accurately capture the scaling with pulse sequence length and the number of π pulses. We outline several strategies to suppress the impact of MW phase noise and implement a version, based on gradiometry, that realizes a >10-fold suppression. Our study highlights an important challenge in the pursuit of sensitive diamond quantum sensors and is applicable to other qubit systems with a large transition frequency.

DOI: 10.1103/PhysRevResearch.6.043148

I. INTRODUCTION

Precision optical measurements of the electron-spin precession of nitrogen-vacancy (NV) centers in diamond form the basis of numerous applications, ranging from imaging biomagnetism [1–6] to nuclear magnetic resonance (NMR) spectroscopy [7–10], gyroscopes [11–13], femtotesla magnetometry [14–17], and searches for new spin physics [18–20]. As solid-state electron-spin sensors, NV centers offer advantages over alkali-metal vapor and superconducting quantum interference device sensors in that a high density of immobile spins form a tunable sensing voxel that can be tailored for the application. Much attention has been devoted to the use of single or small ensembles of NV centers as nanoscale sensors, leading to remarkable advances in materials and biological microscopy [21–24]. However, the development of ultrasensitive bulk NV sensors, as needed for the most demanding metrology applications [25,26], presents additional challenges.

One challenge is that NV electron-spin transition frequencies are almost always large. Due to the zero-field splitting arising from embedding electron-spin $S = 1$ defects in a solid, the magnetic-dipole-allowed NV spin transitions are $\approx$3 GHz at low magnetic field. A femtotesla-level NV sensor must measure changes in these spin transition frequencies at the sub-mHz level, corresponding to a fractional resolution substantially better than 10^{-12}. In any spin-precession measurement, random-phase fluctuations of the microwave (MW) control field lead to undesired rotations of the spin state that are often indistinguishable from magnetic-field noise, see Fig. 1(a). Phase noise is always present at some level, due to the limited clock stability and Johnson noise in MW signal generators [27,28], but it has only recently been a limiting factor for NV precision measurements [9,15,17]. The magnitude of this effect should not be understated–phase noise from typical MW generators produces NV sensor noise at the

*These authors contributed equally to this work.
†Contact author: smitsjanis@gmail.com
‡Contact author: vmacosta@unm.edu

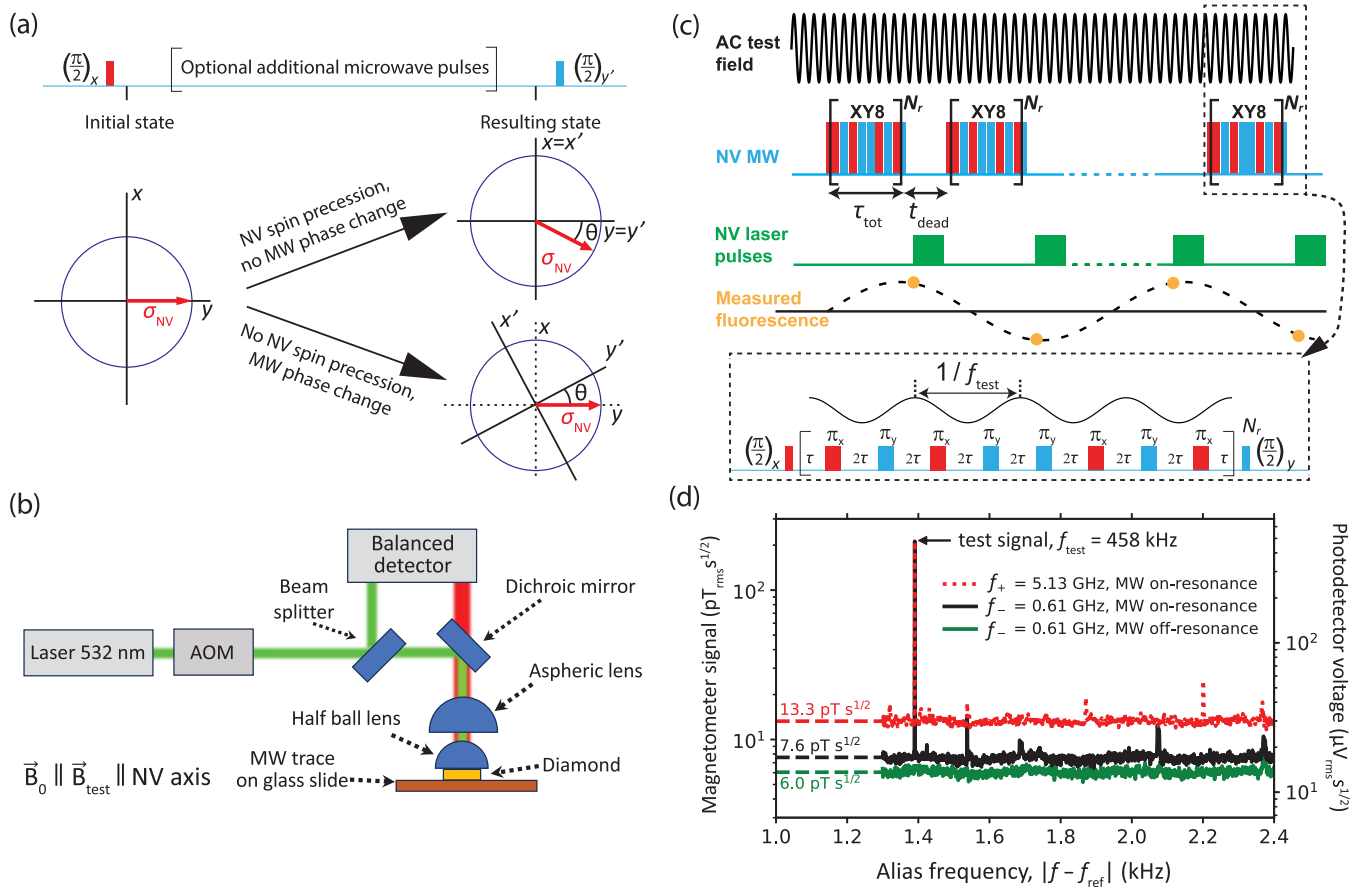

FIG. 1. MW phase-noise detection: principle and experimental apparatus. (a) Graphical representation of the impact of MW phase noise on a generic spin-1/2 sensor. MW phase errors lead to a rotation of the basis, with respect to the spin Bloch vector $\hat{\sigma}_{\rm nv}$, that is indistinguishable from spin precession due to a magnetic field. (b) Schematic of the NV magnetometer used for the measurements; detailed information can be found in Appendix A 4. The bias field $\vec{B}_0$ and calibrated test fields $\vec{B}_{\rm test}$ are aligned to one of the NV axes (Appendix A 3). (c) Synchronized series of XY8-N_r pulse sequences used for AC magnetic-field measurements. (d) Typical magnetic spectra taken using MW generator G2. The right vertical axis is the processed photodetector voltage spectral density, and the left vertical axis is converted to magnetic sensitivity units.

level of 0.1–100 pT s$^{1/2}$, orders of magnitude above requirements for applications like magnetoencephalography [29–33]. Moreover, the impact of phase noise tends to grow with the MW frequency, which has implications for high-field applications like NMR chemical analysis.

Here, we experimentally characterize the impact of MW phase noise on the noise floor of an NV sensor. We show how to predict the impact numerically, using a frequency-domain model incorporating the MW phase-noise spectrum and the filter-function response of the sensing protocol, and also with simple analytic expressions for the cases of white and random-walk phase noise. We discuss strategies to suppress the impact of MW phase noise and implement a version based on gradiometry that realizes a >10-fold suppression. Our results highlight a key factor in the design of sensitive diamond quantum sensors and are broadly applicable to other qubit systems.

II. EXPERIMENTAL SETUP

The principle behind the impact of MW phase noise is depicted in Fig. 1(a). Consider a general Ramsey-type measurement on a spin $S = 1/2$ qubit initially prepared in an S_z eigenstate. A resonant MW $\pi/2$ pulse, with a well-defined phase, rotates the spin state to lie along the y axis of the Bloch sphere in the rotating frame. In the first case, a small additional magnetic field is present, and the spin state precesses, accumulating a phase θ with respect to the y axis. A final MW $\pi/2$ pulse is applied with the same frequency as the initial $\pi/2$ pulse but with a 90° phase shift. This pulse faithfully projects the phase accumulation onto the S_z basis, and measurements of $\langle S_z \rangle$ can be used to determine θ. In the second case, there is no additional magnetic field present, and the spin state does not precess. However, if the phase of the MW carrier of the final $\pi/2$ pulse has an error θ with respect to the desired phase, then the state projection and readout will produce the same result as in the first case. Thus, in this common scenario, MW phase noise is indistinguishable from magnetic-field noise. This general concept holds even when there are additional MW pulses in the sequence, although the exact phase errors depend on the properties of the MW source and the pulse sequence.

Our experimental setup is depicted in Fig. 1(b) and additional details are in Appendix A. A 120-μm-thick, (110)-polished diamond membrane, with an NV density ≈0.5 ppm, is adhered to a high-refractive-index ($n = 2$) half-ball lens.

TABLE I. Microwave generators used in measurements. The final column is the manufacturer specifications of typical phase noise for a 1 GHz carrier at 20 kHz offset with I/Q modulation enabled [34,35]. Other information on phase-noise performance is in Fig. 2(d) and Appendix C.

Name	MW generator model	Phase noise, 1 GHz
G1	SRS SG386	−114 dBc/Hz (20 kHz)
G2	R&S SMATE200A	−134 dBc/Hz (20 kHz)
G3	R&S SMU200A-B22	−138 dBc/Hz (20 kHz)

Light from a 532-nm laser passes through an acousto-optic modulator and is relayed by a 0.79-NA aspheric lens onto the diamond membrane, resulting in an excitation beam waist of 50–100 μm and a power of ≈0.4 W. NV center fluorescence is collected by the same aspheric lens, spectrally filtered (passing 650–800 nm light), and focused onto one channel of a balanced photodetector. A small portion of the green excitation beam is directed to the second photodetector channel to suppress the impact of laser intensity fluctuations. The diamond together with the half-ball is adhered to a microscope slide that has a copper trace which delivers the MW field to the NV centers. A variable bias magnetic field $\vec{B}_0$, produced by an electromagnet, is aligned along one of the NV axes (Appendix A 3). The aligned NV centers have spin transition frequencies $f_\pm = D \pm \gamma_{\rm nv} B_0$, where $D = 2.87$ GHz is the NV zero-field splitting and $\gamma_{\rm nv} = 28.03$ GHz/T is the gyromagnetic ratio. A 55-mm-diameter wire loop is used to deliver uniform, calibrated (Appendix B) oscillating (AC) test magnetic fields along the NV axis. Three commercial MW generators were studied, named G1, G2, and G3. Their model numbers and typical phase-noise performance are shown in Table I.

AC magnetometry is performed at room temperature using a synchronized series of XY8-N_r pulse sequences [8,9,15], see Fig. 1(c). Each XY8-N_r sequence (duration: $\tau_{\rm tot}$) begins and ends with a MW $\pi/2$ pulse that is resonant with one of the $f_\pm$ spin transitions. Between the $\pi/2$ pulses, a train of $8N_r$ resonant π pulses (length: t_π), spaced by $2\tau = 1/(2f_{\rm xy8}) - t_\pi$, are applied in a pattern of alternating phase. The XY8-N_r pulse sequences are frequency selective, in that NV centers are primarily sensitive to AC magnetic-field frequencies within a band, centered at $f_{\rm xy8}$, of width $\approx f_{\rm xy8}/(4N_r)$. Following each XY8-$N_r$ sequence, a 12-μs laser pulse is applied for NV optical readout and repolarization. The time between XY8-N_r sequences, $t_{\rm dead} \approx 15$ μs, accounts for this pulse as well as small additional delays. The sequence is repeated continuously, and the resulting time trace of NV fluorescence readouts is approximately proportional to an aliased version of an AC field that is sampled at the time of the first $\pi/2$ pulse of each XY8-N_r sequence, with a sample rate $f_{\rm samp} = 1/(\tau_{\rm tot} + t_{\rm dead})$. For an AC field of frequency f, that lies within the passband of the XY8-N_r sequence, the NV fluorescence signal oscillates at an alias frequency $|f - f_{\rm ref}|$, where $f_{\rm ref}$ is the integer multiple of $f_{\rm samp}$ that is closest to f [9].

Figure 1(d) shows typical magnetic spectra taken using MW generator G2, with a 212 $\rm pT_{rms}$ test field applied at $f_{\rm test} = f_{\rm xy8} = 457.9$ kHz and $f_{\rm ref} = 459.3$ kHz. Here, $B_0 = 81$ mT, so $f_+ = 5.13$ GHz and $f_- = 0.61$ GHz. To acquire a spectrum, a synchronized XY8-8 pulse series is applied continuously for ≈150 s, the NV fluorescence time trace is split into 1-s intervals, and the root-mean-squared average [36] of the absolute value of the Fourier transform of each interval is computed (Appendix D). Three spectra are shown: one with G2 tuned to the f_+ transition, one with G2 tuned to f_-, and one with G2 detuned by +0.4 GHz from f_- for all MW pulses. The noise floor off resonance is $\eta_{\rm off} \approx 6.0\ \rm pT_{rms}\ s^{1/2}$, consistent with the expected noise floor in the photoelectron-shot-noise limit, $\eta_{\rm psn} \approx 5.4\ \rm pT_{rms}\ s^{1/2}$ (Appendix E). Measurements with G2 detuned by +0.4 GHz from f_+ have the same noise floor and are omitted for visual clarity. However, the spectrum for G2 on resonance with f_- has a slightly higher noise floor, and the spectrum for G2 on resonance with f_+ has a more than twofold higher noise floor. Assuming that the noise contributions are independent, we compute an excess noise as $\eta_{{\rm ex},\pm} = (\eta_\pm^2 - \eta_{\rm off}^2)^{1/2}$, where $\eta_\pm$ is the noise floor when MW are resonant with the $f_\pm$ transitions. The nonzero values, $\{\eta_{{\rm ex},-} \approx 4.7\ \rm pT_{rms}\ s^{1/2}, \eta_{{\rm ex},+} \approx 11.9\ \rm pT_{rms}\ s^{1/2}\}$, indicate additional noise that is only present when G2 is tuned to resonance. This effect is reproducible regardless of the order of acquiring the spectra, the use of spectral filters in the MW chain, the type of MW amplifier used, or the method of phase alternation (Appendix F). However, the results change dramatically when using different MW generators, as will be described in Sec. III. That the noise floor would rise with increasing MW carrier frequency is an early hint that MW phase noise, which also tends to increase with carrier frequency [28], is responsible for the excess noise.

III. PHASE NOISE SPECTROSCOPY

To study the excess noise in more detail and verify its origin, we measured the NV sensor noise floor as a function of MW carrier frequency and detection frequency $f_{\rm xy8}$ for two different MW generators—G1 and G2. Figure 2(a) plots the NV sensor noise floor as a function of G1 carrier frequency $f_{\rm G1}$ under an XY8-6 pulse series. To generate this plot, B_0 is varied over the range 20–80 mT. For a given value of B_0, the noise floor is measured for G1 tuned to an $f_\pm$ resonance, as well as the off-resonance case (detuning: 0.2 GHz), and the excess noise $\eta_{{\rm ex},\pm}$ is inferred. Throughout, we observe markedly larger values of $\eta_{{\rm ex},\pm}$ for G1 than those observed with G2 [see Fig. 1(d)] at the same carrier frequency. This is consistent with the higher phase noise of G1, as specified by the manufacturer (Table I). Furthermore, we observe a roughly monotonic increase in excess noise with increasing MW generator frequency. A similar trend is observed for generator G2 (Appendix A 4) when sweeping B_0 and matching $f_{\rm G2}$ to the $f_\pm$ resonances.

To isolate the MW carrier-frequency dependence, and eliminate any effects due to magnetic field, we next fixed the magnetic field at $B_0 = 81$ mT. We mixed the output of G1 with one of the output channels of (low-noise) G2 and filtered for the difference frequency, $f_{\rm G1} - f_{\rm G2}$. Since G2 has a much lower phase noise (Table I), we assume the phase noise of the difference frequency is dominated by the G1 phase noise (Appendix G). For MW-on-resonance measurements, we set $f_{\rm G1} - f_{\rm G2} = f_- = 611$ MHz, and for

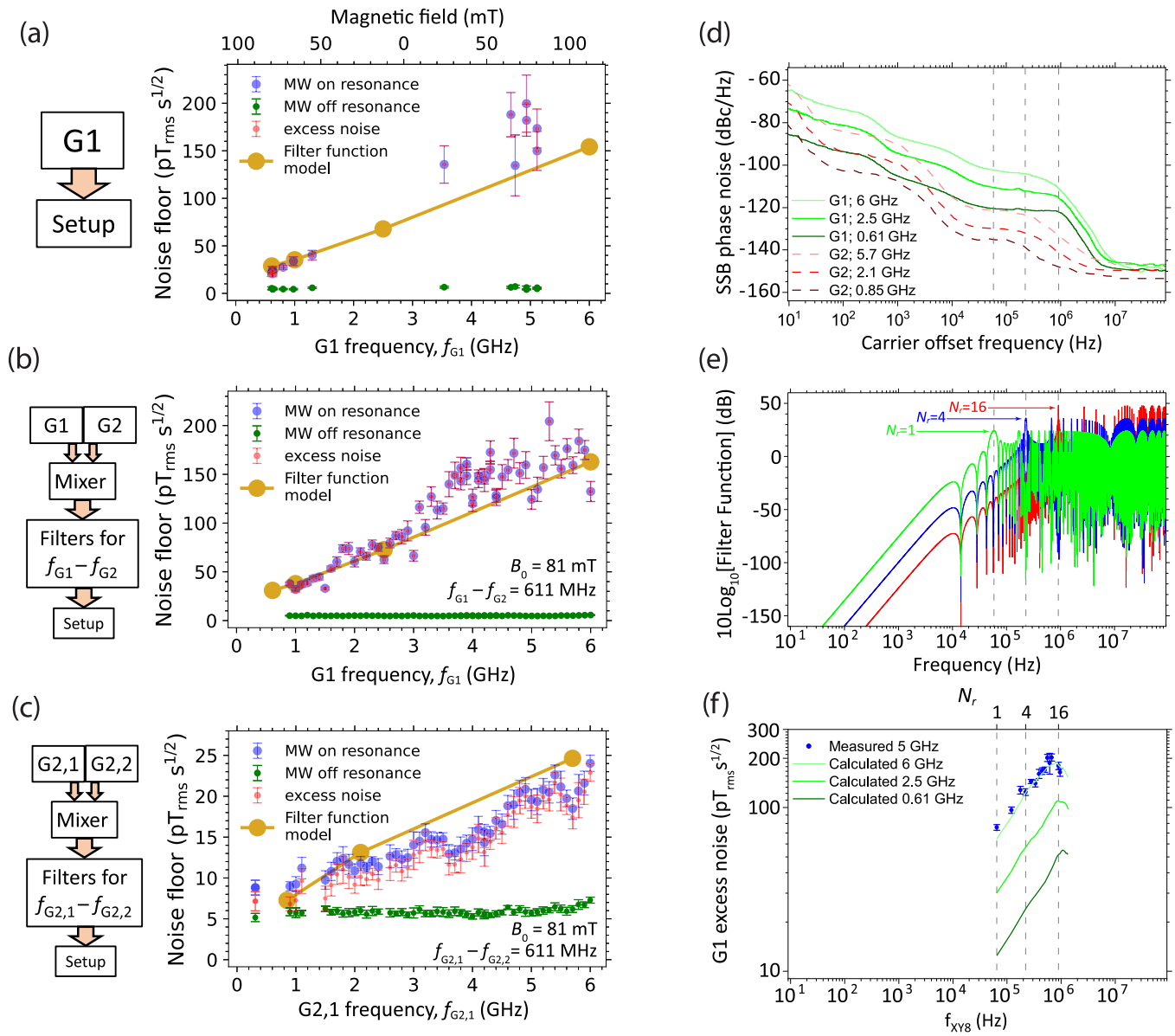

FIG. 2. Phase noise spectroscopy. (a) NV sensor noise floor for an XY8-6 pulse series with MW generator G1 tuned to a $f_\pm$ spin transition at different magnetic fields. For G1 carrier frequency $f_{G1} < 2.87$ GHz, microwaves are tuned to the f_- spin transition, and otherwise they are tuned to f_+. (b) NV sensor noise floor as a function of f_{G1}. Here $B_0 = 81$ mT is constant, and the final MW tone probes the $f_- = 611$ MHz transition. This tone is generated by jointly varying and mixing G1 and G2 carrier frequencies, while maintaining $f_{G1} - f_{G2} = 611$ MHz. (c) NV sensor noise floor as a function of f_{G2}. Here, two independent channels of G2 are varied and mixed, while maintaining $f_{G2,1} - f_{G2,1} = 611$ MHz. The lowest frequency represents $f_{G2,1} + f_{G2,2} = 611$ MHz (Appendix A 4). (d) G1 and G2 single-sideband (SSB) phase-noise power spectra, $\mathcal{L}(f)$, for different MW carrier frequencies. For G1, the spectra are measured using a phase-noise analyzer, Appendix C. For G2, they are extracted from the manufacturer specification sheet [35]. (e) XY8-N_r filter functions, $\mathcal{F}(f)$, for $N_r = \{1, 4, 16\}$. Dashed gray lines are aligned with the first harmonic, f_{XY8}, of each function and are extended to (d) and (f) for comparison. (f) Excess noise, $\eta_{ex,+}$, for G1 obtained for different values of N_r. Here $B_0 = 76$ mT, $f_{G1} = f_+ = 5.00$ GHz, and $\tau_{tot} \approx 70$ µs (Appendixes D and H). Varying N_r also varies the filter function's first harmonic, f_{XY8}. Green curves in panel (f) and gold points in panels (a)–(c) are calculated from the filter-function model using Eqs. (1), (2) and curves in panels (d) and (e). For panels (a)–(c) and (f), error bars are the standard deviation of $\geqslant$10 identical measurements (Appendix D).

the off-resonance case we set $f_{G1} - f_{G2} = 811$ MHz. This allowed us to vary f_{G1} without altering the NV properties. Figure 2(b) shows the measured NV sensor noise floors, under an XY8-8 pulse series, as a function of f_{G1} in the 0.9–6 GHz range. The behavior is similar to that of Fig. 2(a), reinforcing the monotonic dependence of $\eta_{ex,\pm}$ on MW carrier frequency.

We used a similar method to probe the NV sensor noise dependence on f_{G2}, Fig. 2(c). In this case, two output channels of G2 are mixed. We expect that the phase noise of each channel is uncorrelated, and thus the phase noise of the difference frequency is $\approx\sqrt{2}$ times larger than that of a single channel (Appendix G). The measured excess noise increases roughly monotonically with $f_{G2,1}$, but it is $\approx$7 times lower than that

observed in Fig. 2(b). This is qualitatively consistent with the lower phase noise for G2 compared with G1.

To predict the impact of MW generator phase noise on the NV sensor noise floor, we apply a frequency-domain, filter-function method [27]. The method incorporates the single-sideband power spectral density $\mathcal{L}(f)$ for G1 and G2 at different carrier frequencies, as shown in Fig. 2(d) (see also Appendix C). The standard deviation of the NV spin state's phase displacement, σ_ϕ, due to MW phase errors in a single pulse sequence is given by (Appendix H)

$$\sigma_\phi^2 = \int_0^\infty S_\phi(f)\mathcal{F}(f)df \approx \int_0^{f_c} S_\phi(f)\mathcal{F}(f)df, \qquad (1)$$

where $S_\phi(f) = 2 \times 10^{\frac{\mathcal{L}(f)}{10}}$, $\mathcal{F}(f)$ is the filter function of the multipulse sequence, and $f_c \approx 0.1$ GHz is a cutoff frequency that depends on the frequency response of the MW delivery (Appendix H 2). Expressions for $\mathcal{F}(f)$ of an XY8-N_r sequence [37] are given in Appendix H, and the cases of $N_r = 1$, 4, and 16 are shown in Fig. 2(e).

The impact of MW phase-noise results in an equivalent magnetic sensitivity given by

$$\eta_\phi \approx \frac{\sigma_\phi}{4\gamma_{\rm nv}\sqrt{\tau_{\rm tot}}}\sqrt{1 + t_{\rm dead}/\tau_{\rm tot}}. \qquad (2)$$

This noise contribution is uncorrelated with that due to photoelectron shot noise and can thus be directly compared with $\eta_{{\rm ex},\pm}$. In Figs. 2(a)–2(c), the noise floors calculated from η_ϕ [Eqs. (1) and (2)] are shown, using the same pulse sequence parameters as in the experiment (see Appendixes D and H). For higher MW carrier frequencies, the calculated and experimental values differ slightly, perhaps due to 1–2 dB disparities in $\mathcal{L}(f)$. However, overall, the calculated values are largely in agreement with the measurements, indicating that the source of $\eta_{\rm ex}$ is indeed from MW phase noise.

The filter-function model of Eqs. (1) and (2) implies that, for a given MW signal generator, the NV sensor noise floor depends on the choice of pulse sequence. To probe this effect, we performed experiments where we fixed $\tau_{\rm tot} \approx 70$ µs (the experimental $\tau_{\rm tot}$ values varied slightly, but for calculations we used the median value $\tau_{\rm tot} = 70$ µs) and varied N_r (Appendixes D and H). Here $B_0 = 76$ mT, and G1 probes the $f_+ = 5$ GHz transition.

Figure 2(f) shows the resulting measured values of $\eta_{{\rm ex},+}$, along with the noise floors calculated from η_ϕ [Eqs. (1) and (2)] for different G1 carrier frequencies. While a comparison under exactly the same carrier frequency was not possible, the experimental and simulated curves have the same shape. We can qualitatively understand the shape as follows. The G1 phase-noise spectrum $S_\phi(f)$ is approximately flat for carrier offset frequencies between 50–500 kHz (white-noise band), and then it falls sharply as $\approx 1/f^2$ for offset frequencies between 0.5 and 5 MHz (random-walk noise band). When the NV sensor detection frequency falls within the white-noise band ($f_{\rm xy8} \lesssim 500$ kHz, $N_r \lesssim 9$), the scaling of the NV sensor excess noise with N_r is dominated by the scaling of the integral of $\mathcal{F}(f)$ within this band, see Eq. (1). For a sequence of periodic π pulses, including XY8-N_r, the peak heights grow as N_r^2, while the number of harmonics that fall within the band shrinks as $1/N_r$. Thus, σ_ϕ^2 grows approximately as N_r and η_ϕ grows approximately as $\sqrt{N_r}$. However, when the NV sensor detection frequency falls within the random-walk noise band ($f_{\rm xy8} \lesssim 500$ kHz, $N_r \lesssim 9$), the N_r^2 increase in $\mathcal{F}(f)$ peak heights is compensated for by the $\approx 1/f^2$ decrease in $S_\phi(f)$, leading to a plateau in excess noise and even a slight decrease when $N_r \gtrsim 14$.

Based on these results, we conclude (i) noise spectroscopy using a multipulse NV sensor carries information on the MW generator phase-noise spectrum that can be inferred through spectral decomposition methods [38–40], and (ii) the choice of pulse sequence has a large impact on the phase-noise-limited NV sensor noise floor and is thus an important consideration in the design of NV precision experiments.

In Appendix I, we also analyze the phase-noise response of a continuous-wave optically detected magnetic resonance measurement. In this case, the phase-noise-limited equivalent magnetic sensitivity tends to be better than that for multipulse sequences, and it depends on the magnetometer bandwidth. However, the impact of phase noise is still important. For example, for a 1-kHz bandwidth magnetometer, the phase-noise-limited equivalent magnetic sensitivity is ≈ 1.4 pT$_{\rm rms}$ s$^{1/2}$ for G1 operated at $f_{\rm G1} = 2.5$ GHz and ≈ 0.3 pT$_{\rm rms}$ s$^{1/2}$ for G2 operated at $f_{\rm G2} = 2.1$ GHz.

IV. WHITE AND RANDOM-WALK PHASE NOISE

We next detail the response of an NV sensor under the controlled injection of MW phase noise. We focus on the common cases of white and random-walk phase noise, for which we derive simple analytic formulas of their impact on an NV pulsed sensor (Appendix J). Here, it is assumed that the MW carrier phase of a given pulse is well defined, but the relative phase changes from pulse to pulse due to phase noise.

For white phase noise, the standard deviation of the NV spin state's phase displacement due to MW phase errors in a single XY8-N_r pulse sequence is $\sigma_\phi = 2\sigma_{\rm wh}\sqrt{N + 1/4} \approx 2\sigma_{\rm wh}\sqrt{N}$, where $\sigma_{\rm wh}$ is the standard deviation of the MW phase error of each pulse, and $N = 8N_r$ is the total number of π pulses. Using Eq. (2), the equivalent NV sensor magnetic sensitivity is then

$$\eta_{\rm wh} \approx \frac{\sigma_{\rm wh}}{\gamma_{\rm nv}}\sqrt{\frac{f_{\rm xy8}}{2}}\sqrt{1 + t_{\rm dead}/\tau_{\rm tot}}. \qquad (3)$$

We experimentally validated the behavior of the NV sensor noise floor under white MW phase noise, by injecting pseudowhite noise into the phase modulation port of G3 (Appendix A 5). Here, $B_0 = 81$ mT and $f_{\rm G3} = f_- = 0.61$ GHz. Figures 3(a)–3(c) show the NV sensor excess noise as a function of $\sigma_{\rm wh}$, $f_{\rm xy8}$, and $\tau_{\rm tot}$, respectively. In each case the experimental values of $\eta_{\rm ex}$ agree well with those calculated from Eq. (3). Interestingly, in the MW phase-noise limit, the noise floor grows as a function of $f_{\rm xy8}$ (for fixed $\tau_{\rm tot}$). This scaling is quite different from the scaling in the photoelectron-shot-noise limit (Appendix E), where the NV noise floor decreases due to improved contrast when the "dynamical decoupling" property of the multipulse sequence improves the NV coherence time, T_2. Furthermore, the MW-phase-noise-limited noise floor hardly changes as $\tau_{\rm tot}$ increases (for fixed $f_{\rm xy8}$). This is also a departure from the photoelectron-shot-noise

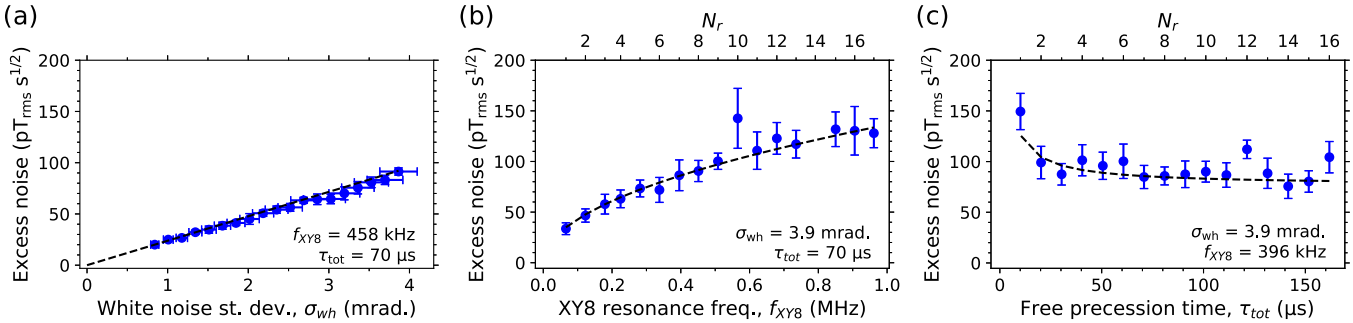

FIG. 3. White phase noise injection. Blue dots are measured values. Black curves are calculated from Eq. (3) using independently measured parameters (Appendix D). (a) NV sensor excess noise as a function of the standard deviation of MW phase errors between π pulses, $\sigma_{\rm wh}$. (b) Excess noise versus N_r (and thus $f_{\rm xy8}$), holding $\tau_{\rm tot}$ and $\sigma_{\rm wh}$ constant. (c) Excess noise versus N_r (and thus $\tau_{\rm tot}$), holding $f_{\rm xy8}$ and $\sigma_{\rm wh}$ constant. In equivalent magnetic noise calculations, we used $t_{\rm dead} = 17$ µs. In panels (a)–(c), error bars are the standard deviation of $\geqslant$10 identical measurements. We used $t_{\rm dead} = 17$ µs for magnetic noise calculations.

limit, where an increase in $\tau_{\rm tot}$ should decrease the noise floor for $\tau_{\rm tot} \lesssim T_2$.

The behavior of the NV sensor noise floor with white phase noise is particularly important, since the white Johnson noise of a MW oscillator sets a fundamental limit on a MW generator's phase-noise performance [28,41]. As discussed in Appendix H, for a 0 dBm MW oscillator with a noise temperature of 300 K, $\mathcal{L} \approx -177$ dBc/Hz, and the phase-noise-limited equivalent magnetic sensitivity would be at the $\approx$100 fT s$^{1/2}$ level. This is far above fundamental limits set by photoelectron shot noise or spin projection noise [25]. In Appendix I, we also analyze the phase-noise response of a continuous-wave optically detected magnetic resonance measurement. We find that the NV sensor equivalent magnetic sensitivity in the oscillator Johnson noise limit is $\gtrsim$10 fT s$^{1/2}$ for a magnetometer bandwidth $\gtrsim$100 kHz.

For random-walk phase noise, the standard deviation of the NV spin state's phase displacement due to MW phase errors in a single XY8-N_r pulse sequence is $\sigma_\phi = \sigma_{\rm rw}\sqrt{\tau_{\rm tot} R_{\rm samp}}$, where $\sigma_{\rm rw}$ is the standard deviation of MW phase jumps and $R_{\rm samp}$ is the jump rate. In deriving this expression (Appendix J), we assumed that $R_{\rm samp} \gtrsim 1/(2\tau)$ so the MW phase changes for each π pulse. The equivalent NV sensor magnetic sensitivity is

$$\eta_{\rm rw} \approx \frac{\sigma_{\rm rw}\sqrt{R_{\rm samp}}}{4\gamma_{\rm nv}}\sqrt{(1 + t_{\rm dead}/\tau_{\rm tot})}. \tag{4}$$

We injected pseudorandom-walk phase noise using G3 (Appendix A 5) and observed the NV noise floor behavior experimentally ($B_0 = 81$ mT, $f_{\rm G3} = f_- = 0.61$ GHz). Figures 4(a)–4(d) show the NV sensor excess noise as a function of $\sigma_{\rm rw}$, $R_{\rm samp}$, $f_{\rm xy8}$, and $\tau_{\rm tot}$, respectively. In each case, the experimental values of $\eta_{\rm ex}$ match those calculated from Eq. (4), aside from minor deviations due to imperfect delivery of random-walk noise in the experiment (Appendix A 5). Unlike the case of white phase noise, with random-walk noise, $\eta_{\rm ex}$ is independent of the number of π pulses or $f_{\rm xy8}$. However, as with white noise (but unlike in the photoelectron-shot-noise limit), $\eta_{\rm ex}$ hardly changes with $\tau_{\rm tot}$. The case of random-walk phase noise is especially relevant for qubit sensors with a long coherence time, as MW generators are often limited by random-walk phase noise at offset frequencies $\lesssim$10 kHz, see, for example, the behavior of G2 in Fig. 2(d).

V. PHASE NOISE CANCELLATION

While MW phase noise represents an important technical challenge for NV precision measurements, various common-mode noise rejection methods can suppress its impact. Here, we implement a simple way to minimize the impact of MW phase noise using two-point gradiometry. The excitation laser beam is split into two beams that are focused to separate spots in the same diamond, located on opposite sides of the MW

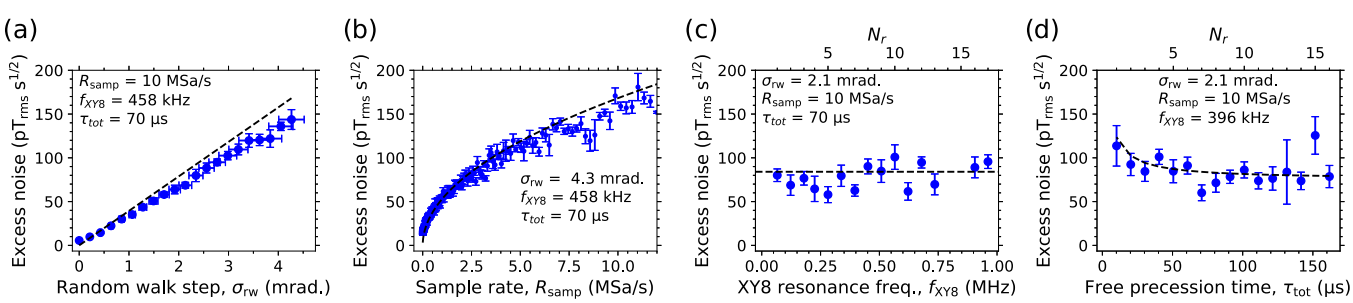

FIG. 4. Random-walk phase-noise injection. Blue dots are measured values. Black curves are calculated from Eq. (4) using independently measured parameters (Appendix D). (a) NV sensor excess noise versus the standard deviation of random-walk phase jumps $\sigma_{\rm rw}$. (b) Excess noise versus jump sample rate, $R_{\rm samp}$. (c) Excess noise versus N_r (and thus $f_{\rm xy8}$), holding $\tau_{\rm tot}$, $R_{\rm samp}$, and $\sigma_{\rm rw}$ constant. (d) Excess noise versus N_r (and thus $\tau_{\rm tot}$), holding $f_{\rm xy8}$, $R_{\rm samp}$, and $\sigma_{\rm rw}$ constant. In panels (a)–(c), vertical error bars are the standard deviation of $\geqslant$5 identical measurements, and we used $t_{\rm dead} = 17$ µs for magnetic noise calculations.

trace (Appendixes A 2 and A 3). The emission from each spot is directed to separate channels of the balanced photodetector. The detection spots are close enough apart (≈0.2 mm) that NV centers in each spot are subject to approximately the same MW field (aside from an overall ≈180° phase difference), and thus similar phase noise. In these experiments, the diamond is rotated within the electromagnet to interrogate one of the out-of-plane NV axes.

Magnetometer spectra were obtained using G1 with a XY8-7 pulse series under two applied test field configurations. In the first configuration, Fig. 5(a), a 394 kHz test field is generated from a large coil (Appendix A), resulting in a uniform test magnetic field over the two spots. The magnetic spectra of each spot separately (emission from the other channel was blocked) feature a 6.3 nT_{rms} test-signal peak, and the MW-phase-noise-limited noise floor is ≈175 $pT_{rms}\,s^{1/2}$. However when operating as a gradiometer (neither channel is blocked), the test signal is highly suppressed to $\lesssim$0.2 nT_{rms}. Moreover, the noise floor drops to 15 $pT_{rms}\,s^{1/2}$, a factor of ≈12 suppression compared with the single-spot case.

In the second configuration, a 394 kHz gradient test field is applied by sending the test signal through the same trace on the chip as the MW radiation, Fig. 5(b). Here, magnetic spectra of each spot separately exhibit a ≈7.5 nT_{rms} test-signal peak, with a MW-phase-noise-limited noise floor of ≈170 $pT_{rms}\,s^{1/2}$. However, when operating as a gradiometer (neither channel is blocked), the test signal peak is approximately doubled to 15 nT_{rms}, which is expected since the sign of the test field is opposite between the two spots. Nevertheless, the gradiometer noise floor drops to 16 $pT_{rms}\,s^{1/2}$, an ≈11-fold suppression compared with the single-spot case.

We tested the stability of the gradiometer as a function of acquisition time t over the course of 5000 s. The noise floor exhibits a 14 $pT_{rms}\,s^{1/2}/\sqrt{t}$ scaling for the duration of the measurement, dropping below 200 fT_{rms} for $t = 5000$ s. Moreover, the gradient signal peak remains at 14.7 ± 0.5 nT_{rms} over the course of the measurement, indicating a stable scale factor.

VI. DISCUSSION

Two-point gradiometry is a powerful method for phase-noise suppression when detecting localized fields that vary substantially over millimeter length scales. In cases where fields vary less sharply, or when only a single detection spot is desired, the use of multiple NV transitions can be exploited.

For example, for a small applied field along the NV axis ($B_0 \ll D/\gamma_{nv} = 0.1$ T), a double-quantum pulse sequence can be applied to suppress the impact of MW phase noise (Appendix K). Two MW tones at $f_\pm$ can be generated by mixing a large carrier frequency ($f_{car} = D$) with a much lower frequency local oscillator ($f_{LO} = \gamma_{nv} B_0$) [42]. As derived in Appendix K, the resulting NV sensor signal then depends only on the phase noise of the local oscillator, and it is independent of the phase noise of the f_{car} source. If the phase noise varies linearly with carrier frequency, as is approximately the case for the MW generators studied here, this scheme allows for a suppression of the MW phase-noise impact by a factor of f_{car}/f_{LO}.

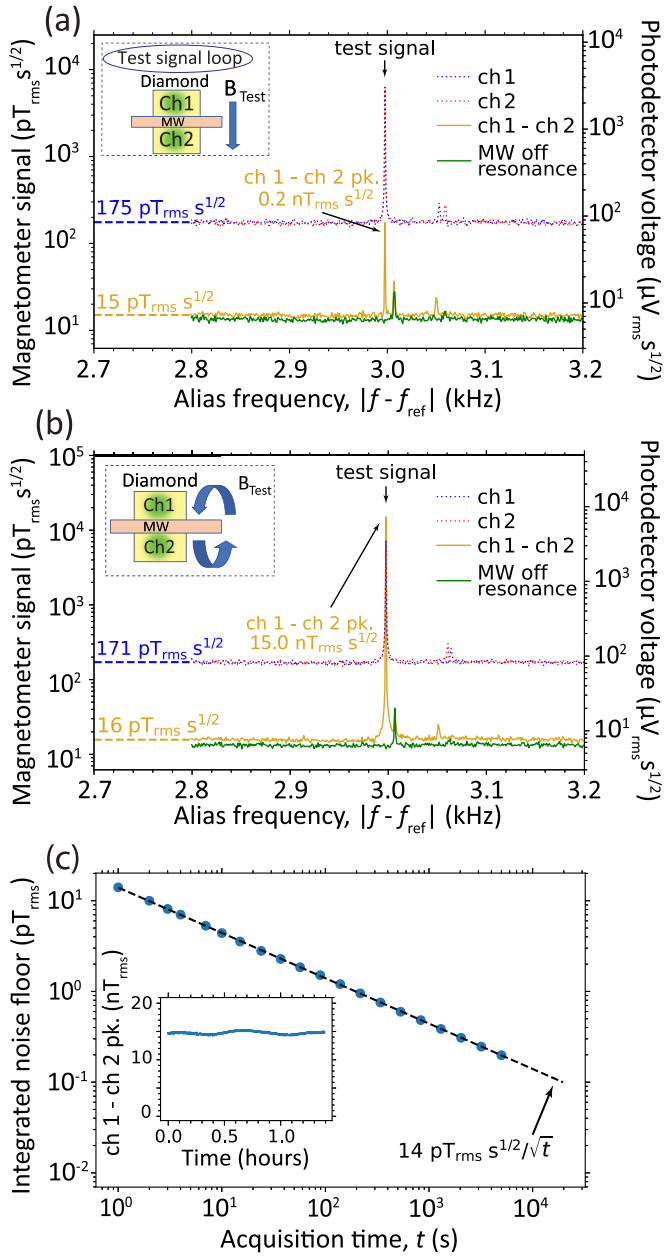

FIG. 5. Phase noise cancellation by gradiometry. G1 is used with a XY8-7 series at $B_0 = 76$ mT, $f_+ = 5$ GHz, $f_{test} = 394$ kHz (Appendix D). (a) Magnetometer spectra obtained with a uniform test signal. Blue and red curves are spectra for individual channels (ch 1 and ch 2), corresponding to emission from each spot separately. The gradiometer difference signal is in yellow. The MW-off-resonance noise for the gradiometer is in green. (b) Magnetometer spectra obtained for a gradient test signal. The gradiometer signal shows a peak that is nearly twice that of individual channels. The noise floor is suppressed by a factor of ≈11 and is near the MW-off-resonance limit. (c) Gradiometer integrated noise floor as a function of acquisition time t. The data are well fit by a 14 $pT_{rms}\,s^{1/2}/\sqrt{t}$ curve out to $t = 5000$ s. (inset) Stability of the gradiometer test signal amplitude over 5000 s.

For a continuous-wave NV measurement, at low magnetic field, the same mixer scheme can be used with a dual-resonance technique [14,43] to realize comparable levels of suppression of the MW phase impact (Appendix I 3).

At very low applied field $B_0 \lesssim \Omega_R/\gamma_{\rm nv}$, where Ω_R is the MW Rabi frequency, a single MW tone can be used with a double-quantum pulse sequence [44], potentially eliminating the first-order impact of MW phase noise entirely.

In some applications, interrogating both NV resonances may not be feasible or effective. In either case, pulse sequences may be improved by incorporating MW phase noise into the optimization of pulse-sequence filter functions. The dependence of the NV phase-noise-limited sensitivity on the filter-function detection frequency, Fig. 2(f), points to the possible efficacy of this strategy. Pulse sequences that take advantage of low-phase-noise regions of a generator's spectrum (for example, due to the behavior of the phase-locked loop), or the use of composite, chirped, or aperiodic pulses, may offer superior performance. When practical, ultralow phase noise MW generators based on superconducting resonators [45], dielectric cavities [46,47], ferrimagnetic oscillators [48,49], and photonic microwave generators [50–52] may be helpful. Finally, the use of magnetic-flux concentrators [14–16,53] provides some relief from the impact of phase noise, as magnetic signals are amplified while the impact of phase noise is unchanged.

In summary, we studied the impact of MW phase noise on NV sensors and verified that it is an important limiting factor for high-sensitivity experiments. We showed how to quantitatively predict the impact of MW phase noise, with knowledge of a generator's phase-noise spectrum and the pulse sequence filter function. We provided simple analytic expressions for the case of white and random-walk phase noise. Finally, we identified a number of methods to suppress the impact of phase noise and implemented one, based on two-point gradiometry, that provides a >10-fold suppression. Our results inform on the design of precision measurements and spectroscopy experiments using NV centers and other qubit systems with a large transition frequency.

ACKNOWLEDGMENTS

We gratefully acknowledge advice and support from P. Schwindt, D. Thrasher, T. Drake, C. Ramanathan, I. Savukov, D. Budker, P. R. Hemmer, A. Jeronimo Perez, and anonymous reviewers. We especially thank A. Orozco for loaning and assisting with measurements using a MW phase-noise analyzer. This work was supported by the National Science Foundation (Grants No. CHE-1945148 and No. OIA-1921199) and the National Institutes of Health (Grants No. DP2GM140921 and No. R41GM145129).

J.S., A.B., M.S.Z., and V.M.A. conceived the idea and designed the experiments. A.B. and M.S.Z. built the main experimental apparatus, acquired and analyzed data, and wrote the initial paper draft. J.T.D., I.F., Y.S., and J.S. made early measurements using previous setups. J.F.B., P.K., A.J., Y.S., B.A.R., and J.S. assisted in theoretical modeling, experimental execution, and data interpretation. M.S.Z. and J.S., wrote the control software. V.M.A. supervised the project. All authors helped edit the paper.

The authors declare no competing financial interests.

APPENDIX A: EXPERIMENTAL SETUP DETAILS

Our diamond sensor is similar to those used in Refs. [9,15]. The diamond was grown by chemical vapor deposition, yielding an initial nitrogen density of ≈5 ppm. The diamond was irradiated with 2-MeV electrons at a dose of 3×10^{17} cm^{-2} and was subsequently annealed at 800 and 1100 °C using the recipe described in Ref. [54]. Based on fluorescence brightness, we estimate the resulting NV$^-$ concentration is ≈0.5 ppm. Finally, the diamond was cut into membranes of dimensions $\approx 800 \times 600 \times 120$ μm^3 with a (110) surface polish.

The diamond sensor is sandwiched between a glass slide with a copper trace and a half-ball lens (8 mm diameter, refractive index $n = 2$). A Norland Optical Adhesive 88 is used at the interfaces to mechanically secure the diamond in place and improve optical contact. Microwave fields are delivered to the NV centers through the ≈2-μm-thick, lithographically defined copper trace, which has a trace width of ≈200 μm in the region in contact with the diamond. The laser light (532 nm, Lighthouse Photonics Sprout-G-10W) passes through an acousto-optic modulator (Brimrose TEM-110-25-532) to form 12 μs pulses with a peak power ≈0.4 W. The light subsequently passes through an additional lens [not shown in Fig. 1(a) of the main text] and is relayed onto the diamond using a NA = 0.79 aspheric lens. The arrangement of lenses (including the $n = 2$ half-ball lens) is selected to provide wide-field illumination with an excitation diameter in the range 50–100 μm. NV fluorescence is collected by the same lens, is spectrally filtered by a dichroic mirror (Thorlabs DMLP550L) and long-pass filter (Thorlabs FELH0650), and is directed to one channel of a balanced photodetector (Thorlabs PDB210A) with a transimpedance gain of $G = 175$ kV/A $= 2.8\times10^{-14}$ V/(photoelectron/s) at 50 Ω impedance.

Considering the ≈50 μm excitation beam diameter, the ≈120 μm diamond membrane thickness, and the ≈0.5 ppm NV density, we estimate that $\approx 2 \times 10^{10}$ NV centers are optically interrogated. Of these NV centers, ≈1/4 ($\approx 5 \times 10^9$) belong to the NV-axis subensemble with $f_\pm$ spin resonances addressed by the microwaves.

An arbitrary waveform generator (Teledyne LeCroy T3AFG80) is used to generate AC test magnetic fields via an external loop surrounding the diamond sensor. The loop is formed from copper magnet wire (diameter 1.3 mm) that is wound three times in a circle of diameter 55 mm. In Fig. 5(a), we used a modified loop that had two windings and an elliptical shape with ≈55-mm-long major axis and ≈20-mm-long minor axis.

A transistor-transistor logic (TTL) pulse card (SpinCore PBESR-PRO-500) is used to generate and synchronize pulse sequences and a data acquisition card (National Instruments USB-3631) is used to digitize the photodetector signal.

1. Microwave components

The following microwave components were used (part numbers are for Mini-Circuits unless otherwise noted):

When addressing the f_- transition using MW carrier frequencies below 2.5 GHz [Figs. 1(d) and 2(a)–2(c)], we used

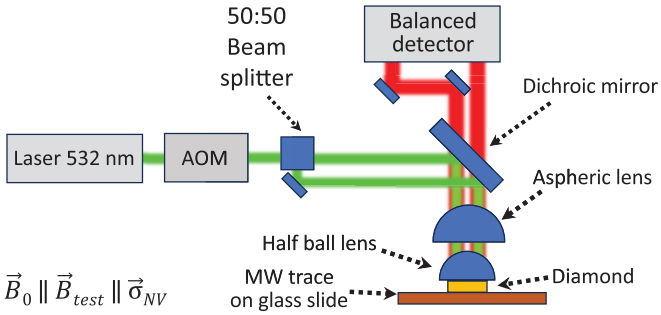

FIG. 6. Schematic of the gradiometry setup. The primary difference from Fig. 1(b) is that the excitation beam is split and focused to two spots on the diamond. The fluorescence from each spot is sent to different channels of the balanced photodetector.

ZHL-30W-252-S+ as the power amplifier. Otherwise, we used ZVE-3W-83+ followed by RF-Lambda RP02G06GSPA.

For experiments presented in Figs. 2(b) and 2(c), we used ZX05-42MH-S+ as the frequency mixer when both MW carrier frequencies were $\lesssim$4 GHz and ZMX-7GHR otherwise. In both cases, we used a combination of SHP-700+ and VLFX-450+ on the mixer output as an effective bandpass filter. To compensate for losses in the mixer and filters, we used an additional pre-amplifier before the power amplifier: ZX60-53LNB-S+ for MW carrier frequency $\lesssim$5 GHz, and either ZVA-1S3WA-S+ or ZVA-183WA-S+ otherwise.

We conducted a number of control experiments to check that the pre-amplifiers, power amplifiers, mixers, and frequency filters did not add additional phase noise. We did not observe additional NV sensor excess noise when swapping out different components. For example, in Fig. 2(b), there are two measured values for the $f_{G1} = 3.9$–4.2 GHz frequency range. One set of values corresponds to the use of the lower-frequency MW components, and the other used the higher-frequency components. No systematic difference in NV sensor excess noise is observed. A similar control is presented in Fig. 2(c) for $f_{G2,1} = 4.1$ GHz.

For the experiment presented in Fig. 5(b), we used a diplexer Marki Microwave DPXN-0R5 to combine the microwave field and test signal.

In all experiments, the microwave power was adjusted to produce a π pulse length of around 40–60 ns, depending on the carrier frequency. The normalized peak-to-peak amplitude of the processed photodetector signal during Rabi oscillations is typically $\approx$6%.

2. Gradiometry phase-noise cancellation

For the two-point gradiometry experiments presented in Fig. 5, we used a modified experimental setup, Fig. 6. The excitation laser beam is split into two beams, which are focused to separate spots on the diamond. The fluorescence from each spot is collected and sent to different channels of the balanced photodetector ("ch 1", "ch 2"). The fluorescence spots are close enough together ($\approx$0.2 mm) that they experience approximately the same microwave field, meaning that both channels also have the same MW phase properties. By replacing the laser light with fluorescence from ch 2 on the balanced photodetector, we still compensate for laser intensity fluctuations, as both channels respond to the same laser intensity fluctuations. This scheme allows us to subtract the impact of MW phase noise by measuring the difference between the two channels of the balanced photodiode.

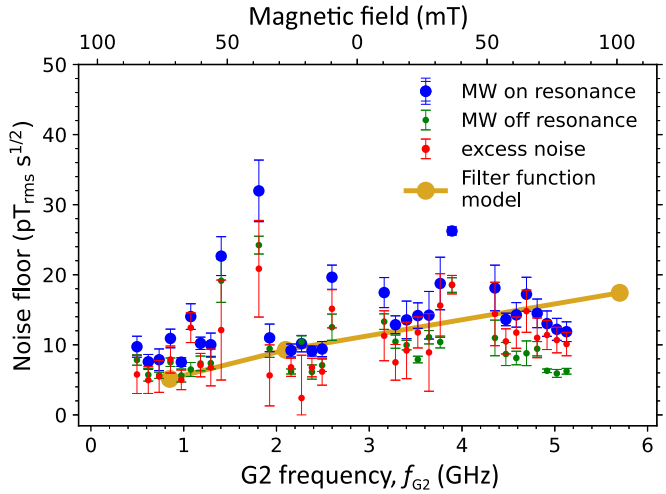

FIG. 7. NV sensor noise floor for MW generator G2 tuned to a $f_\pm$ spin transition at different magnetic fields. For G2 carrier frequency $f_{G2} < 2.87$ GHz, microwaves are tuned to the f_- spin transition; otherwise they are tuned to f_+.

For these experiments, we used one of the out-of-diamond-plane NV axes. This was chosen because, in our setup, a test signal applied through the MW trace [Fig. 5(b)] produces a field that is orthogonal to the in-plane NV axes. Producing a test field with a significant component along the NV axis was only possible for the two out-of-plane NV axes. To align $\vec{B}_0$ with an out-of-plane NV axis, we rotated the diamond in the electromagnet and modified the optical beam path accordingly.

3. Magnetic-field geometry

The bias magnetic field $\vec{B}_0$ is produced by an electromagnet, and the direction of $\vec{B}_0$ is aligned along one of the NV axes by rotating the diamond. For all measurements in Figs. 1–4, $\vec{B}_0$ is aligned with an NV axis in the plane of the diamond faces. For two-point gradiometry, Fig. 5, the magnetic field is aligned with an out-of-plane NV axis.

For almost all the measurements, the test field is also aligned along one of the NV axes. The exception is the experiment presented in Fig. 5(b), where a gradient test field is applied through the MW trace. The gradient test field is approximately normal to the surface (at the location of the fluorescence spots) and thus makes a $\approx 35^\circ$ angle with respect to the relevant out-of-plane NV axis for our (110) diamond.

4. Specifics of phase-noise spectroscopy measurements

In Fig. 2(a), we presented the NV sensor excess noise as a function of G1 MW carrier frequency. The $f_\pm$ resonances were swept by tuning B_0. Figure 7 shows a similar measurement using G2 under an XY8-6 pulse series. As in Fig. 2(a), to generate this plot, B_0 is varied in the range 20–80 mT. For a given value of B_0, the noise floor is measured for f_{G2} tuned

to a $f_\pm$ resonance, as well as the off-resonance case, and the excess noise $\eta_{\mathrm{ex},\pm}$ is inferred. In Fig. 7, the data are noisy, but there are hints of the same general trend of an increase in NV sensor noise with MW carrier frequency.

For the mixer experiments in Figs. 2(b) and 2(c), we hold $B_0 = 80.7$ mT constant. A frequency mixer is used to produce a MW signal at the difference frequency between two inputs. For Fig. 2(b), we use G1 and G2 as inputs, and for Fig. 2(c) we use two channels of G2. The source generating the higher-frequency input is the one that is I/Q phase modulated in the XY8-N_r pulse series. The difference frequency is always set to be 611 MHz. This configuration allows us to vary the MW generator frequency without having to vary B_0 or the NV transition frequencies. We can thus eliminate any effects due to the frequency response of different MW components, as well as the NV sensor's dependence on magnetic field (e.g., the vicinity of $B_0 = 0$, fields corresponding to NV cross-relaxations with other electron spins [55], and fields corresponding to ^{13}C-induced NV echo envelope collapses [56]).

In Fig. 2(c), the measurement point at the lowest frequency was obtained by using the mixer to create the frequency sum ($f_{\mathrm{G2,1}} + f_{\mathrm{G2,2}} = 310.5$ MHz $+ 300.0$ MHz $=$ 610.5 MHz). The reason why we did this is because the lowest-frequency pair we were able to use to generate a clean 611 MHz difference signal was approximately $\{0.3, 0.911\}$ GHz. The additional sum-frequency point allowed us to extend our observations to lower frequencies. Note that the phase noise of the mixer output is always assumed to be the quadrature sum of the phase noise of the individual inputs, see Appendix G.

5. Specifics of injected phase noise

For the noise injection experiments of Figs. 3 and 4, voltage noise was generated using an arbitrary waveform generator and sent to the phase modulation port of G3. The random walk waveform has $\approx 1.6\times 10^6$ points and the waveform is repeated continuously in a loop. Each waveform is the concatenation of 1000 random-walk sequences. Each sequence has a variable length of mean 1600 points and standard deviation 100 points. The variable length is chosen to suppress any signatures of the sequence length in the frequency domain. The choice of a 1600-point mean sequence length was made to ensure that voltage extrema in any sequence did not exceed the maximum input voltage of the phase modulation port. Each phase jump (the difference between adjacent points) is drawn from a normal distribution with a standard deviation σ_{rw} specified for each experiment. Varying σ_{rw} is realized by varying the G3 phase modulation port's scale factor. The first and last points of the waveform are set to zero to accommodate the arbitrary waveform generator resetting behavior.

The white-noise waveform comprises 1.6 million points drawn from a zero-mean normal distribution. The first and last points of the waveform are also set to zero, and the waveform repeats in a continuous loop.

The noise generated in this manner is not truly random. To verify that our injected noise protocol is still a good approximation to random-walk or white noise, we conducted experiments with different waveforms that were generated in the same manner. The results were consistent.

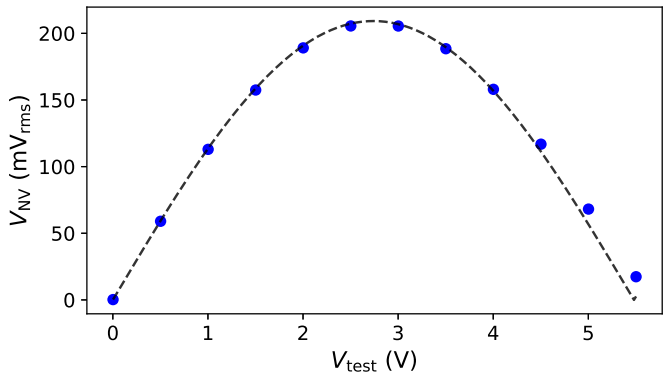

FIG. 8. Plot of the NV sensor test-signal peak V_{nv} as a function of the voltage amplitude of the test signal, V_{test}. A fit to Eq. (B1) reveals κ, the voltage-to-field conversion factor.

In all phase-noise-injection experiments, a 11 MHz low-pass filter was used on the arbitrary waveform generator output. The phase modulation port of G3 also has a pass band of DC to $\approx$10 MHz. The limited bandwidth had some implications on the effective values of σ_{rw} and σ_{wh}. We independently measured the frequency response by mixing two channels of G3, one of which was phase modulated with the noise waveforms. From analysis of the homodyne signal, we were able to directly measure σ_{wh}. We found that this value was $\approx$20% lower than the expected value if the noise injection had infinite bandwidth. In all measurements in the paper, we report the slightly lower value that was directly measured. For σ_{rw}, we find the limited bandwidth also leads to attenuation, but the exact value depends on the sampling interval. However, for most experiments in the text, the sampling interval is consistent with a $\approx$15% attenuation, so we use this value throughout.

APPENDIX B: CALIBRATED TEST FIELDS

To calibrate the AC test signals, we used a method similar to that of Refs. [8,15]. A voltage signal, with amplitude V_{test} and frequency f_{test}, is applied to the test loop, producing a homogenous AC test field within the diamond. The amplitude V_{test} is varied and the NV signal is recorded at each value of V_{test}. The magnitude of the NV fluorescence photodetector signal due to the test field, V_{nv}, is given by

$$V_{\mathrm{nv}} = V_{\mathrm{max}} |\sin(4\sqrt{2}\kappa V_{\mathrm{test}}\gamma_{\mathrm{nv}}\tau_{\mathrm{tot}})|, \tag{B1}$$

where V_{max} is the maximum NV signal magnitude, $\kappa = B_{\mathrm{test}}/V_{\mathrm{test}}$ is a fitted scaling factor, and $\tau_{\mathrm{tot}} = 8(2\tau + t_\pi)N_{\mathrm{r}}$ is the interval between $\pi/2$ pulses in the XY8-N_r sequence (see Sec. II of the main text). The fitted calibration factor, κ, allows us to convert the applied test signal amplitude to magnetic-field units.

Figure 8 shows an example plot of $V_{\mathrm{nv}}(V_{\mathrm{test}})$ along with a fit to Eq. (B1) used to calibrate the test signal in Fig. 1(d). We typically repeat the calibration process before and/or after acquiring a dataset to account for drifts of the test loop position or NV collection region.

APPENDIX C: INDEPENDENT MICROWAVE PHASE-NOISE MEASUREMENTS

The phase-noise spectra $\mathcal{L}(f)$ of the G1 and G3 MW generators were measured directly with a Berkeley Nucleonics 7000 Phase Noise Tester. Further direct G1 phase noise spectra were acquired with a Rohde & Schwarz FSWP26 [see Fig. 2(d)]. The G1$\mathcal{L}(f)$ curves obtained with the two instruments were similar and also comparable to the manufacturer's specifications. We did not measure $\mathcal{L}(f)$ curves of G2 directly. However, we found that the directly measured $\mathcal{L}(f)$ curves for G3 (a similar model from the same manufacturer) matched closely those reported as "typical" by the manufacturer. This lends confidence that our use of the manufacturer-specified $\mathcal{L}(f)$ curves for G2 is a decent approximation.

All direct $\mathcal{L}(f)$ measurements were performed in a continuous-wave mode, without pulsing, but with I/Q modulation enabled. When the MW power input to the amplifier was low ($\lesssim -10$ dBm), the single-sideband (SSB) phase-noise values were higher than specifications at higher offset frequencies ($\gg$1 kHz); this effect is negligible at the high input powers used in our experiments for the frequency bands of interest. When adding a MW switch, and/or using a power amplifier and attenuators, there was no significant effect on the measured $\mathcal{L}(f)$ curves aside from minor variations at offset frequencies $>$10 MHz.

APPENDIX D: DATA ACQUISITION AND PROCESSING

NV sensor noise signals were acquired in two modes. In the first mode, used for most experiments including Figs. 2(a)–2(c), 2(f), 3, and 4, a synchronized XY8-N_r pulse series of length $\approx$30 ms is applied, and the test signal is synchronized with the start of the series using the burst mode of a function generator. We repeat the pulse series $N_{\mathrm{avg}} \approx 35$ times, and the photodetector time trace of each series is averaged in the time domain. The absolute value of the Fourier transform of the averaged data is computed, and the noise floor is extracted as described in Appendix D. The above process takes $\approx$1 s, and it is repeated for $\gtrsim$5 runs to gather statistics. The median and standard deviation of the noise floors is computed, serving as, respectively, the point estimate and uncertainty in the figures.

The second mode is used for data presented in Figs. 1(d) and 5. Here, a synchronized XY8-N_r pulse series is applied continuously for $\approx$150 s. The photodetector time trace is split into 1-s intervals, and the absolute-value-squared of the Fourier transform (power spectral density) of each interval is computed. The square root of the mean of the power spectral density is obtained, producing a final "smoothed" spectrum. In Sec. II of the main text, we call this spectrum "the root-mean-squared average [36] of the absolute value of the Fourier transform of each interval."

Experimental parameters used to generate Figs. 1, 2, and 5 are shown in Table II. Experimental parameters used to generate Figs. 3 and 4 are shown in Table III. Some parameters used in the analytic calculations in these figures are the median of experimental values, as indicated in their insets.

TABLE II. Experimental parameters used in the measurements presented in Figs. 1, 2, and 5.

Fig.	N_r	t_π (ns)	τ (ns)	τ_{tot} (μs)	t_{dead} (μs)	f_{xy8} (kHz)
1(d)	8	48	522	69.9	15.0	458
2(a)	6	48	622	62.0	14.9	387
2(b)	8	48	522	69.9	14.7	458
2(c)	8	48	522	69.9	14.7	458
2(f)	1	58	3844	62.0	27.5	64.5
2(f)	2	58	2046	66.4	23.1	120
2(f)	3	58	1388	68.1	21.4	176
2(f)	4	58	1100	72.3	17.1	221
2(f)	5	58	872	72.2	17.2	277
2(f)	6	58	720	72.0	17.3	333
2(f)	7	58	612	71.9	17.4	389
2(f)	8	58	532	71.9	17.5	445
2(f)	9	58	468	71.7	17.5	502
2(f)	10	58	418	71.7	17.5	558
2(f)	11	58	378	71.8	17.6	613
2(f)	12	58	344	71.8	17.6	668
2(f)	16	58	250	71.7	17.7	893
2(f)	17	58	232	71.3	17.6	954
5(a)–5(c)	7	46	612	71.1	17.1	394

Noise floor estimation

The NV sensor noise floor is computed as the median of the absolute value of the Fourier transform in "spike-free" regions. The noise floor can be verified visually, as seen in Fig. 1(d). For experiments requiring the estimation of multiple noise floors, the process of defining spike-free regions is automated. First, the frequency range of 0–1 kHz and a $\approx$1 kHz band about the test signal frequency are excluded. Next, the median and standard deviation of the remaining spectrum, excluding the highest 10% of spectral points, is computed. Spectral points that are more than four standard deviations larger than the median are identified as spikes and excluded. The noise floor is then given by the median of the remaining spectrum.

This automated process to calculate the noise floor was tuned and routinely checked for accuracy based on visual inspection of the spectrum. In some datasets, this procedure is replaced by visually locating a spike-free region near the test signal frequency and computing the median of the noise floor in this region.

APPENDIX E: PHOTOELECTRON-SHOT -NOISE-LIMITED SENSITIVITY

The minimum detectable magnetic field due to photoelectron shot noise for a single XY8-N_r readout is given by [15,25]

$$\Delta B_{\mathrm{single}} \approx \frac{\xi}{4\gamma_{\mathrm{nv}} C \tau_{\mathrm{tot}} \sqrt{N_{\mathrm{ph}}}}. \tag{E1}$$

Here $\xi = \sqrt{2(1 + t_{\mathrm{r}}/t_{\mathrm{n}})}$ accounts for the extra photoelectron noise due to the balanced detection and normalization procedure, where t_{r} is the readout duration at the start of the laser pulse, t_{n} is the duration of normalization at the end of the laser pulse, C is the effective contrast of the XY8-N_r sequence, and

TABLE III. Experimental parameters used for measurements presented in Figs. 4 and 3.

Fig.	N_r	t_π (ns)	τ (ns)	$\tau_{\rm tot}$ (μs)	$t_{\rm dead}$ (μs)	$f_{\rm xy8}$ (kHz)
4(a)	8	48	522	69.9	14.7	458
4(b), 3(a)	8	48	522	69.9	14.7	458
4(c), 3(b)	1	44	3768	60.6	27.7	66.0
4(c), 3(b)	2	44	2015	65.2	23.0	124
4(c), 3(b)	3	44	1372	66.9	21.3	179
4(c), 3(b)	4	44	1090	71.2	17.0	225
4(c), 3(b)	5	44	866	71.0	17.1	282
4(c), 3(b)	6	44	718	71.0	17.2	338
4(c), 3(b)	7	44	610	70.8	17.3	396
4(c), 3(b)	8	44	532	70.9	17.3	451
4(c), 3(b)	9	44	470	70.9	17.3	508
4(c), 3(b)	10	44	420	70.7	17.4	566
4(c), 3(b)	11	44	380	70.8	17.4	622
4(c), 3(b)	12	44	346	70.7	17.4	679
4(c), 3(b)	13	44	318	70.7	17.4	735
4(c), 3(b)	16	44	254	70.7	17.5	906
4(c), 3(b)	17	44	238	70.7	17.5	962
4(d), 3(c)	1	44	610	10.1	17.6	396
4(d), 3(c)	2	44	610	20.2	17.5	396
4(d), 3(c)	3	44	610	30.3	17.5	396
4(d), 3(c)	4	44	610	40.5	17.4	396
4(d), 3(c)	5	44	610	50.6	17.4	396
4(d), 3(c)	6	44	610	60.7	17.3	396
4(d), 3(c)	7	44	610	70.8	17.3	396
4(d), 3(c)	8	44	610	80.9	17.2	396
4(d), 3(c)	9	44	610	91.0	17.2	396
4(d), 3(c)	10	44	610	101.1	17.1	396
4(d), 3(c)	11	44	610	111.2	17.1	396
4(d), 3(c)	12	44	610	121.3	17.0	396
4(d), 3(c)	13	44	610	131.5	17.0	396
4(d), 3(c)	14	44	610	141.6	16.9	396
4(d), 3(c)	15	44	610	151.7	16.9	396
4(d), 3(c)	16	44	610	161.8	19.3	396

$N_{\rm ph}$ is the number of photoelectrons detected in the readout phase.

The theoretical limit for the magnetometer sensitivity due to photoelectron shot noise is defined as

$$\eta_{\rm psn} = \Delta B_{\rm single}\sqrt{\tau_{\rm tot}+t_{\rm dead}} \approx \frac{\xi}{\sqrt{\delta}}\frac{1}{4\gamma_{\rm nv}C\sqrt{\tau_{\rm tot}N_{\rm ph}}}, \quad \text{(E2)}$$

where $\delta = \tau_{\rm tot}/(\tau_{\rm tot}+t_{\rm dead})$ is the measurement duty cycle. Using the experimental parameters of Fig. 1(d) ($t_{\rm n} = 4$ μs, $t_{\rm r} = 1.5$ μs, $\tau_{\rm tot} = 69.9$ μs, $t_{\rm dead} = 15$ μs, $C = 0.013$, $N_{\rm ph} = 1.23\times10^9$), we estimate $\eta_{\rm psn} \approx 4.3$ pT$_{\rm rms}$ s$^{1/2}$. To convert to a noise floor in the absolute value of the Fourier transform, this quantity is multiplied by a factor $\sqrt{\pi/2} \approx 1.253$, resulting in $\eta_{\rm psn} \approx 5.4$ pT$_{\rm rms}$ s$^{1/2}$. The latter is the value that is directly comparable to our experimental noise-floor measurements.

In Eqs. (E1) and (E2), the estimation of $N_{\rm ph}$ is made by recording the peak voltage of the photodetector fluorescence-channel's monitor port. While this level varied somewhat from experiment to experiment, for the data in Fig. 1(d) it was $V_{\rm mon} \approx 0.39$ V at 50 Ω impedance. Taking into account the difference in gain between the two ports inferred from the manufacturer specifications, this is equivalent to an effective voltage in the RF difference port (at 50 Ω impedance) of $V_0 \approx 23$ V. We could not use the RF port to measure V_0 directly because it saturates at ≈3.5 V. The estimation of C is obtained by taking the peak value $V_{\rm max} = 0.2$ V$_{\rm rms}$ in Fig. 8. From Eq. (B1) (see also Ref. [15]), the contrast is given by $C \approx V_{\rm max}\sqrt{2}/V_0 \approx 0.013$.

APPENDIX F: EARLY ATTEMPTS TO UNDERSTAND THE EXCESS NOISE

As mentioned in the main text, when we first observed the presence of excess noise, $\eta_{{\rm ex},\pm}$, we did several tests to confirm its nature. Things that did not change the noise behavior included

(1) swapping amplifiers. Instead of an RF-Lambda RP02G06GSPA, we used a Mini-Circuits ZHL-25W-63+. A small increase in the contrast C of the magnetometer was observed, but $\eta_{{\rm ex},\pm}$ was unchanged. We believe the improvement in C is due to the higher output power of the Mini-Circuits amplifier improving the microwave pulse fidelity.

(2) modifying our method of I/Q modulation. Instead of using switches to apply DC voltages to the I/Q modulation ports of the signal generators, we used a single switch toggling between two out-of-phase 300 MHz signals from a function generator. No perceptible change in noise floor was observed.

(3) reducing the MW Rabi frequency, increasing the pulse lengths. We observed a small reduction in C, but $\eta_{{\rm ex},\pm}$ was unchanged. Further reduction in MW Rabi frequency degraded the sensitivity to the point where the $\eta_{{\rm ex},\pm}$ could no longer be distinguished from the photoelectron shot-noise.

(4) changing the lengths of the cables carrying TTL signals to the I/Q modulation ports of the signal generators. Cable lengths were adjusted such that the phase toggling for the MW pulses happened exactly between any two pulses firing. We observed a small improvement in C, but $\eta_{{\rm ex},\pm}$ was unchanged.

APPENDIX G: PHASE NOISE WHEN MIXING TWO FREQUENCIES

As discussed in Sec. III of the main text, a mixer was used to probe the impact of phase noise on the NV sensor noise floor as a function of microwave carrier frequency. For pedagogical reasons, we derive the well-known operation of an ideal mixer here.

If a mixer is assumed to be perfect, two input signals $v_1(t) = v_{01}\sin(\omega_1 t+\alpha_1)$ and $v_2(t) = v_{02}\sin(\omega_2 t+\alpha_2)$ in the input would yield an output signal:

$$v_3 = v_{03}\sin(\omega_1 t+\alpha_1)\sin(\omega_2 t+\alpha_2). \quad \text{(G1)}$$

By applying the product-to-sum identity the above expression can be rewritten as

$$v_3 = \frac{v_{03}}{2}\{\cos[(\omega_2-\omega_1)t+(\alpha_2-\alpha_1)] + \cos[(\omega_2+\omega_1)t+(\alpha_2+\alpha_1)]\}. \quad \text{(G2)}$$

In the experiments depicted in Figs. 2(b) and 2(c) a filter is used to keep only the difference-frequency term. The exception is the lowest-frequency point of Fig. 2(c), where only the sum-frequency term is kept. In either case, applying the

filtering process to Eq. (G2) leaves

$$\tilde{v}_3 = \frac{v_{03}}{2}\cos(\Delta\omega t + \Delta\alpha), \tag{G3}$$

where $\Delta\omega = \omega_2 \pm \omega_1$ and $\Delta\alpha = \alpha_2 \pm \alpha_1$.

In the presence of phase noise, the α_1 and α_2 phases of the signals fluctuate in time. If the two signals are generated independently, then α_1 and α_2 are uncorrelated, and the standard deviation of the output phase $\Delta\alpha$ is

$$\sigma_{\Delta\alpha} = \sqrt{\sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2}, \tag{G4}$$

where σ_{α_1} and σ_{α_2} are the standard deviations of the input phases, α_1 and α_1, respectively.

From equation (G4) it can be seen that if $\sigma_{\alpha_1} \gg \sigma_{\alpha_2}$ then $\sigma_{\Delta\alpha} \approx \sigma_{\alpha_1}$. This is the approximation applied for the measurements in Fig. 1(b), where the phase noise of G1 is much higher than that of G2. If $\sigma_{\alpha_1} \approx \sigma_{\alpha_2}$ then $\sigma_{\Delta\alpha} \approx \sqrt{2}\sigma_{\alpha_1}$. This is the approximation applied for the measurements in Fig. 2(c), where two independent channels of the same generator, G2, are used. The assumption that $\sigma_{\alpha_1} \approx \sigma_{\alpha_2}$ starts to break down if the ratio of the carrier frequencies ω_1/ω_2 deviates substantially from unity. In Fig. 2(c), this is the region $f_{G2,1} \lesssim 1$ GHz where we did not acquire data using the difference frequency. However, the approximation still holds when using the sum frequency, as long as the two carrier frequencies are comparable, as was the case for our lowest-frequency measurement in Fig. 2(c). We thus were able to apply the approximation uniformly to all the data in Fig. 2(c).

APPENDIX H: FREQUENCY-DOMAIN MODEL USING FILTER-FUNCTION APPROACH

The phase-noise-limited equivalent magnetic sensitivity can be calculated for arbitrary phase-noise character, as long as the MW phase-noise spectrum is known. For this, we use the independently measured phase-noise spectrum of the MW signal generator [see Fig. 2(d)] and the frequency-domain filter function of the pulse sequence (identical for the Carr-Purcell-Meiboom-Gill [CPMG] sequence or XY8-N_r). The single-sideband power spectral density of the local oscillator's phase fluctuations in a MW signal generator, $S_\phi(f)$ (in units of rad^2/Hz), can be written as [28]

$$S_\phi(f) = 2 \times 10^{\frac{\mathcal{L}(f)}{10}}, \tag{H1}$$

where $\mathcal{L}(f)$ is the single-sideband power spectral density of the phase noise expressed in logarithmic units of dBc/Hz, and f is the carrier offset frequency.

The standard deviation of the NV spin state's phase displacement σ_ϕ due to MW phase errors is given by

$$\sigma_\phi^2 = \int_0^\infty S_\phi(f)\mathcal{F}(f)df \approx \int_0^{f_c} S_\phi(f)\mathcal{F}(f)df, \tag{H2}$$

where $\mathcal{F}(f)$ is the frequency-domain filter function of the measurement sequence and f_c is a cutoff frequency above which $S_\phi(f)\mathcal{F}(f) \approx 0$. In practice, the value of f_c is influenced by low-pass filtering of MW phase noise due to the MW circuit spectral response, any bandpass filters in the MW chain, and dissipation from the various MW components and coaxial cables (Appendix H 2). The filter function for a pulse sequence with equally spaced π pulses (including CPMG and XY8-N_r) is given by [37]

$$\mathcal{F}(f) = \left| 1 + (-1)^{N+1} e^{i2\pi f \tau_{\text{tot}}} + 2\sum_{j=1}^{N} (-1)^j e^{i\frac{j-1/2}{N}2\pi f \tau_{\text{tot}}} G(t_\pi) \right|^2, \tag{H3}$$

where $G(t_\pi)$ is a term that modifies the filter function to take into account the finite length of the π pulses, t_π. In the limit $t_\pi \to 0$, the NV total phase accumulation time is $\tau_{\text{tot}} = 2N\tau$ and $G(t_\pi) = 1$. If t_π cannot be neglected, we take $\tau_{\text{tot}} = N(2\tau + t_\pi)$ and $G(t_\pi) = \cos(\pi f t_\pi)$ [37].

TABLE IV. Parameters used to calculate η_ϕ, using Eqs. (H2), (H3), and (H5). The resulting η_ϕ values are shown in Tables V and VI and plotted in Fig. 2.

Fig.	N_r	t_π (ns)	τ (ns)	τ_{tot} (μs)	t_{dead} (μs)	f_{xy8} (kHz)
2(a)	6	48	622	62.0	14.85	390
2(b)	8	48	522	69.9	14.85	461
2(c)	8	48	522	69.9	14.85	461
2(f)	–	60	–	71.9	17.1	–

TABLE V. Calculated phase-noise-limited equivalent magnetic sensitivity values for G1, using the filter function model, as presented in Figs. 2(a) and 2(b). * denotes $\mathcal{L}(f)$ was measured independently using a phase-noise analyzer. "spec" denotes $\mathcal{L}(f)$ was extracted from the manufacturer's specifications sheet [34], shown for comparison purposes. The values of η_{ex} shown in Fig. 2 used finite pulse duration ($t_\pi = 48$ ns) and the measured $\mathcal{L}(f)$ spectra. In Fig. 2(a) and 2(b), each of the η_ϕ values in this table are multiplied by a factor of $\sqrt{\pi/2} = 1.253$ to compare with the experimental noise floor, η_{ex}. Here, the maximum frequency offset is taken to be $f_c = 0.1$ GHz.

Fig.	f_{G1} (GHz)	t_π (ns)	η_ϕ (pT$_{\text{rms}}$ s$^{1/2}$)
2(a)	0.61*	0	24.9
2(a)	0.61*	48	23.0
2(a)	1.00 (spec)	0	30.2
2(a)	1.00 (spec)	48	27.5
2(a)	1.00*	48	28.2
2(a)	2.50*	0	56.0
2(a)	2.50*	48	54.1
2(a)	6.00 (spec)	0	107.0
2(a)	6.00 (spec)	48	104.9
2(a)	6.00*	48	123.0
2(b)	0.61*	0	26.7
2(b)	0.61*	48	24.7
2(b)	1.00 (spec)	0	32.5
2(b)	1.00 (spec)	48	29.7
2(b)	1.00*	0	32.4
2(b)	1.00*	48	30.2
2(b)	2.50*	0	60.1
2(b)	2.50*	48	58.0
2(b)	6.00 (spec)	0	113.5
2(b)	6.00 (spec)	48	111.2
2(b)	6.00*	48	129.9

TABLE VI. Calculated phase-noise-limited sensitivity values values for G2, using the filter function model, as presented in Fig. 2(c). In all cases, $\mathcal{L}(f)$ was extracted from the manufacturer's specifications sheet [35]. The values of $\eta_{\rm ex}$ shown in Fig. 2(c) used finite pulse duration ($t_\pi = 48$ ns). In Fig. 2(c), each of the η_ϕ values in this table are first multiplied by a factor of $\sqrt{2}$ to account for the mixer addition (Appendix G) and then multiplied by a factor of $\sqrt{\pi/2} = 1.253$ to compare with the experimental noise floor, $\eta_{\rm ex}$. Here, the maximum frequency offset is taken to be $f_c = 0.1$ GHz.

Fig.	$f_{\rm G2}$ (GHz)	t_π (ns)	η_ϕ (pT$_{\rm rms}$ s$^{1/2}$)
2(c)	0.85	0	5.4
2(c)	0.85	48	4.1
2(c)	2.10	0	9.4
2(c)	2.10	48	7.4
2(c)	5.70	0	15.9
2(c)	5.70	48	13.9

Fluctuations in the phase of the NV spin state, Eq. (H2), result in an equivalent magnetic noise. For a single readout of a pulse sequence, the equivalent magnetic noise standard deviation is given as [25]

$$\Delta B_{\rm single} \approx \frac{\sigma_\phi}{4\gamma_{\rm nv}\tau_{\rm tot}}, \tag{H4}$$

where $\gamma_{\rm nv}$ is the NV gyromagnetic ratio and $\tau_{\rm tot}$ is the NV total phase accumulation (or sensing) time. When averaging multiple-pulse sequences, each of duration $\tau_{\rm tot} + t_{\rm dead}$, for a total acquisition time T, there are $T/(\tau_{\rm tot} + t_{\rm dead})$ independent measurements. In this case, the equivalent magnetic noise standard deviation becomes

$$\Delta B_\phi = \frac{\Delta B_{\rm single}}{\sqrt{T/(\tau_{\rm tot} + t_{\rm dead})}}. \tag{H5}$$

Using Eqs. (H4) and (H5), the MW-phase-noise limited equivalent magnetic sensitivity is given by

$$\eta_\phi = \Delta B_\phi \sqrt{T} \approx \frac{\sigma_\phi}{4\gamma_{\rm nv}\sqrt{\tau_{\rm tot}}}\sqrt{1 + t_{\rm dead}/\tau_{\rm tot}}. \tag{H6}$$

The NV sensor noise floor calculations shown in Fig. 2 were computed using Eqs. (H6), (H2), and (H3). The parameters used in the calculations are shown in Table IV. The resulting calculated values of $\eta_{\rm ex}$ are listed in Tables V and VI, using two approximations for $G(t_\pi)$ in Eq. (H3). In the first, we assume infinitesimal pulses $t_\pi = 0$ and thus $G = 1$. In the second, we assume finite pulses ($t_\pi = 48$ ns) and $G(t_\pi) = \cos(\pi f t_\pi)$. We use the finite-pulse approximation for values in Fig. 2 of the main text, but the differences are small. Note that the magnetic noise floor of the absolute value of a Fourier transform ($\eta_{\rm ex}$) is defined slightly differently from the equivalent magnetic sensitivity (η_ϕ) in Eq. (H6). To convert, we take $\eta_{\rm ex} = \eta_\phi\sqrt{\pi/2} \approx 1.253\eta_\phi$.

1. Oscillator Johnson phase-noise limit: Pulsed

A fundamental source of phase noise in any practical MW generator is that due to Johnson noise, which is frequency-independent (white) and only depends on temperature. If the local oscillator of a signal generator is driven by 1 mW (0 dBm) power near room temperature, the minimum MW phase noise of the output carrier in 1 Hz bandwidth is $\mathcal{L}(f) = \mathcal{L}_{\rm min} \approx -177$ dBc/Hz [28,41].

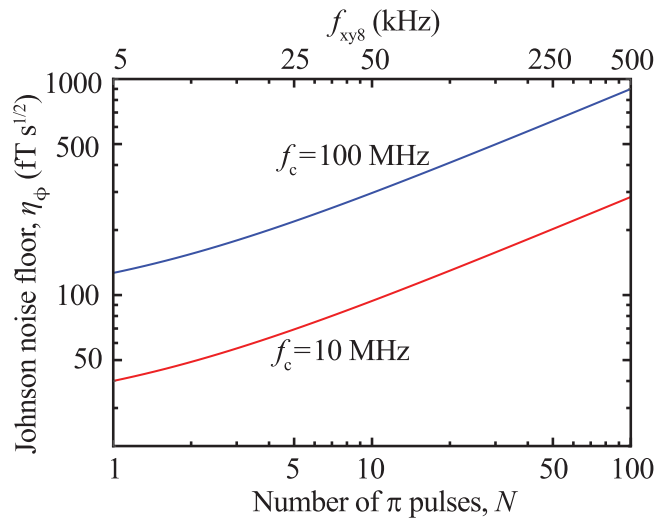

FIG. 9. Oscillator Johnson phase-noise limit: pulsed. Equivalent magnetic sensitivity in the oscillator-Johnson-phase-noise limit for a CPMG or XY8 multipulse sequence, $\eta_{\phi,J}(N)$, calculated using Eq. (H9), with $\tau_{\rm tot} = 50$ µs and $f_c = 10$ MHz (red) or 100 MHz (blue). The first harmonic of the pulse sequence's filter function, $f_{\rm xy8} = N/(2\tau_{\rm tot})$ is shown in the top horizontal axis.

The variance in the NV spin's phase displacement, Eq. (H2), due to the MW Johnson phase errors can be written as

$$\sigma^2_{\phi,J} \approx 2 \times 10^{\frac{\mathcal{L}_{\rm min}}{10}} \int_0^{f_c} \mathcal{F}(f)df. \tag{H7}$$

For the multipulse filter function in Eq. (H3), the integral on the right-hand side of Eq. (H7) is independent of the pulse length τ_π and spacing between the pulses (τ) if $f_c \gtrsim 1/(2\tau_\pi)$, and it can be approximated as [27]

$$\int_0^{f_c} \mathcal{F}(f)df \approx \begin{cases} (4N+2) \times 2\pi f_c, & \tau_\pi \to 0 \\ (2N+2) \times 2\pi f_c, & \tau_\pi \neq 0. \end{cases} \tag{H8}$$

Using Eqs. (H6), (H7), and (H8), the equivalent magnetic sensitivity due to oscillator Johnson phase noise in a multipulse NV sensor with $\tau_\pi \neq 0$ and $\tau_{\rm tot} \gg t_{\rm dead}$ is

$$\eta_{\phi,J} \approx \frac{1}{\gamma_{\rm nv}}\sqrt{\frac{\pi f_c(N+1) \times 10^{\mathcal{L}_{\rm min}/10}}{2\tau_{\rm tot}}}. \tag{H9}$$

Figure 9 shows $\eta_{\phi,J}$ as a function of the number N of π pulses calculated using Eq. (H9), with $\tau_{\rm tot} = 50$ µs and $f_c =$ 10 MHz or 100 MHz.

2. Comment on choice of cutoff frequencies

The use of a hard cutoff frequency f_c for integrals in the filter function model, for example Eq. (1) of the main text, is an approximation. For calculations in Fig. 2 in the main text, we select $f_c = 0.1$ GHz. We found that the exact choice of f_c did not have much impact on our estimates of η_ϕ in Fig. 2(a)–2(c), 2(f), as long as we chose $f_c \gtrsim 10$ MHz. This can be attributed to the strong roll-off of $\mathcal{L}(f)$ at offset

frequencies $f \gtrsim 1$ MHz for the generators studied here, see Fig. 2(d).

For calculations of the equivalent magnetic sensitivity in the oscillator Johnson phase-noise limit in a pulsed experiment, Appendix H 1, the results depend strongly on the choice of f_c, as $\eta_{\phi,J} \propto \sqrt{f_c}$, see Eq. (H9). In Fig. 9, we plot $\eta_{\phi,J}(N)$ for two feasible values $f_c = 10$ MHz and $f_c = 100$ MHz. These cutoff frequencies could be due to the limited passband realized with a MW cavity. Lower values of f_c may also be realized with even higher-quality-factor MW cavities, but this may come with additional challenges in pulse fidelity as the MW passband becomes comparable to the NV spin-resonance spectral width.

A similar trend is found in the calculations for a continuous-wave magnetometer presented in Appendix I. The estimations of η_{f} based on $\mathcal{L}(f)$ curves of G1 and G2 depend only weakly on the choice of f_c (at least for $f_c \gtrsim 10$ kHz). However, for calculations of the equivalent magnetic sensitivity in the oscillator Johnson phase-noise limit the choice is important, as $\eta_{f,J} \propto \sqrt{f_c}$, see Eq. (I5). In Appendix I 1, we plot $\eta_{f,J}(\tau)$ for $f_c = 1$ MHz. This choice of f_c may be justified by the assumed NV response to frequency-detuning fluctuations, but a more accurate value of f_c will depend on specific experimental conditions.

3. Comment on the use of magnetic-flux concentrators

In the main text, we point out that the use of magnetic-flux concentrators offers some relief from the impact of phase noise for magnetometry at appropriately low bias fields. The reason is that magnetic signals are amplified while the impact of microwave phase noise is unchanged. A magnetic-flux concentrator is a ferromagnetic material that collects magnetic flux from a larger area and concentrates it into a microscale diamond sensor [14–16,53]. The enhancement factor ϵ is the ratio of the concentrated field inside the diamond to the unperturbed external field, $\epsilon = B_{\mathrm{int}}/B_{\mathrm{ext}}$. When limited by noise sources fundamental to the diamond detector (such as photon shot noise or phase noise), the sensitivity of a flux-concentrator diamond sensor to external magnetic fields is given by $\eta_{\mathrm{ext}} = \eta_{\mathrm{int}}/\epsilon$, where η_{int} is the magnetic sensitivity of the diamond without concentrators.

In Ref. [15], a similar diamond and pulse sequence as those in this work were used, in addition to a magnetic-flux concentrator that provided $\epsilon \approx 300$. Also, the same generator G3 was used at carrier frequencies of 2.7-3 GHz. The internal magnetic sensitivity was measured to be in the range $\eta_{\mathrm{int}} = 18$–$22\ \mathrm{pT\,s^{1/2}}$ for detection frequencies in the range $f_{\mathrm{xy8}} = 0.2$–1.5 GHz. The equivalent noise floor when MW were detuned was slightly lower, more consistent with the photon-shot-noise-limited prediction of $\approx 12\ \mathrm{pT\,s^{1/2}}$. Using the filter-function model of Eqs. (1) and (2) for G3 under these pulse sequences, the equivalent magnetic sensitivity due to phase noise is at the level of ≈ 5–$10\ \mathrm{pT\,s^{1/2}}$. This level of noise is consistent with the quadrature difference between on-resonance and off-resonance noise floors observed in Ref. [15]. However, with flux concentrators, the overall sensor sensitivity is expected to be $\eta_{\mathrm{ext}} = \eta_{\mathrm{int}}/\epsilon \approx 20\ \mathrm{pT\,s^{1/2}}/300 = 70\ \mathrm{fT\,s^{1/2}}$, which is comparable to what was observed. Thus, Ref. [15] provides evidence that, when the diamond sensor is limited by phase noise, the use of magnetic-flux concentrators can still improve magnetic sensitivity down to the femtotesla level.

APPENDIX I: PHASE NOISE IN A CONTINUOUS -WAVE EXPERIMENT

It is instructive to briefly look at how phase noise limits the sensitivity in a continuous-wave (cw) NV magnetometer, where both the green laser and microwave radiation are applied continuously. The microwave frequency is chosen such that a small change in spin-resonance frequency corresponds to the maximum, linear change in the fluorescence rate. Assuming a Lorentzian lineshape, the fluorescence as a function of MW frequency detuning, $\Delta f(t)$, can be written as

$$F(t) \approx F_0\left(1 - C\frac{3\sqrt{3}}{4}\frac{\Delta f(t)}{\Gamma}\right), \tag{I1}$$

where F_0 is the fluorescence level, C is the resonance contrast, and Γ is the full-width-at-half-maximum resonance linewidth. Here, $\Delta f(t)$ can be caused either by a small magnetic field or by a time-dependent fluctuation in frequency due to phase noise, $\Delta f(t) = \gamma_{\mathrm{nv}} B + \delta f(t)$.

The frequency fluctuations $\delta f(t)$ are assumed to be zero mean, with a variance that depends on the sampling interval τ, given by [28]

$$\sigma_f^2(\tau) \approx \int_0^{f_c} f^2 S_\phi(f) \frac{\sin^2(\pi f \tau)}{(\pi f \tau)^2}. \tag{I2}$$

Here we have approximated the NV response as having a hard high-frequency cutoff at f_c, where the value of f_c is influenced by experimental parameters such as laser intensity, microwave field strength, and the NV spin coherence times and level dynamics. The final (sinc^2) term in the integrand of Eq. (I2) is the filter function of the operation of averaging $\delta f(t)$ over the sampling interval τ.

Since MW frequency fluctuations affect the fluorescence signal in the same way as magnetic-field noise, we can define an equivalent magnetic noise standard deviation for a single measurement over a sampling interval τ as

$$\sigma_B(\tau) = \frac{\sigma_f(\tau)}{\gamma_{\mathrm{nv}}}. \tag{I3}$$

Using Eq. (I3), we define a sampling-interval-dependent equivalent magnetic sensitivity as

$$\eta_f(\tau) = \sigma_B(\tau)\sqrt{\tau}. \tag{I4}$$

We can use Eqs. (I2), (I3), and (I4), along with a known MW generator phase-noise spectrum $S_\phi(f)$, to estimate $\sigma_B(\tau)$ and $\eta_f(\tau)$ in a cw experiment. Figure 10(a) shows the sampling rate dependence of $\sigma_B(\tau)$ for generators G1 and G2 using the phase-noise spectra shown in Fig. 2(d) for $f_{\mathrm{G1}} = 2.5$ GHz, $f_{\mathrm{G2}} = 2.1$ GHz, and $f_c = 1$ MHz. Figure 10(b) shows $\eta_f(\tau)$ as a function of the maximum detectable frequency (i.e., the magnetometer bandwidth), $1/(2\tau)$. In the cw case, the phase-noise-limited equivalent magnetic sensitivity tends to be lower than that for multipulse sequences, and it depends on the magnetometer bandwidth. However, the impact of phase noise is still important. For example, for a 1-kHz

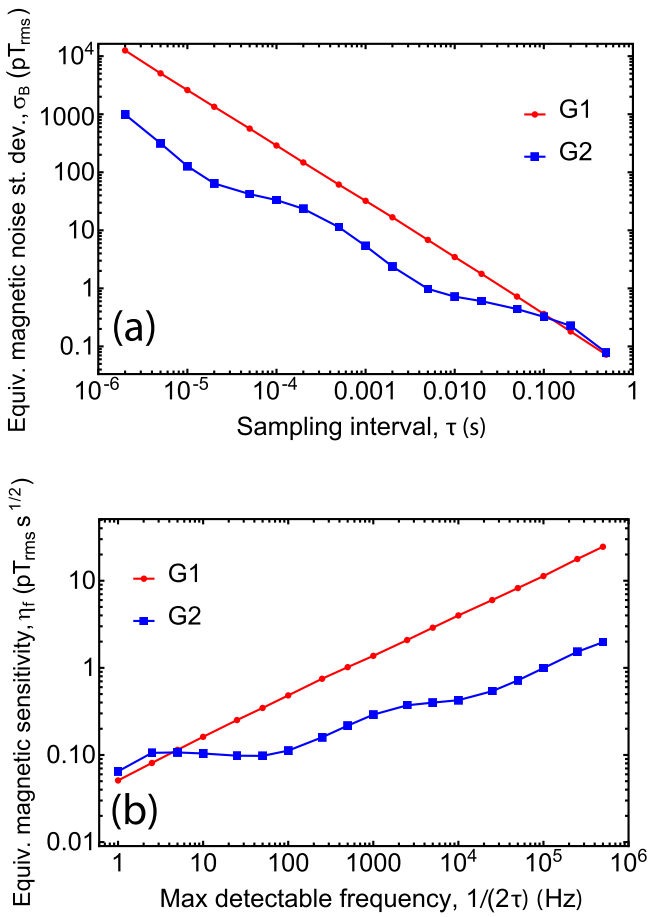

FIG. 10. Magnetic noise in a cw experiment. (a) Estimates for equivalent magnetic noise standard deviation σ_B as a function of sampling interval τ, for generators G1 and G2 at $f_{\mathrm{G1}} = 2.5$ GHz and $f_{\mathrm{G2}} = 2.1$ GHz, respectively [see Fig. 2(d) for the $\mathcal{L}(f)$ curves used]. (b) Estimated equivalent magnetic sensitivity η_f as a function of magnetometer bandwidth $1/(2\tau)$.

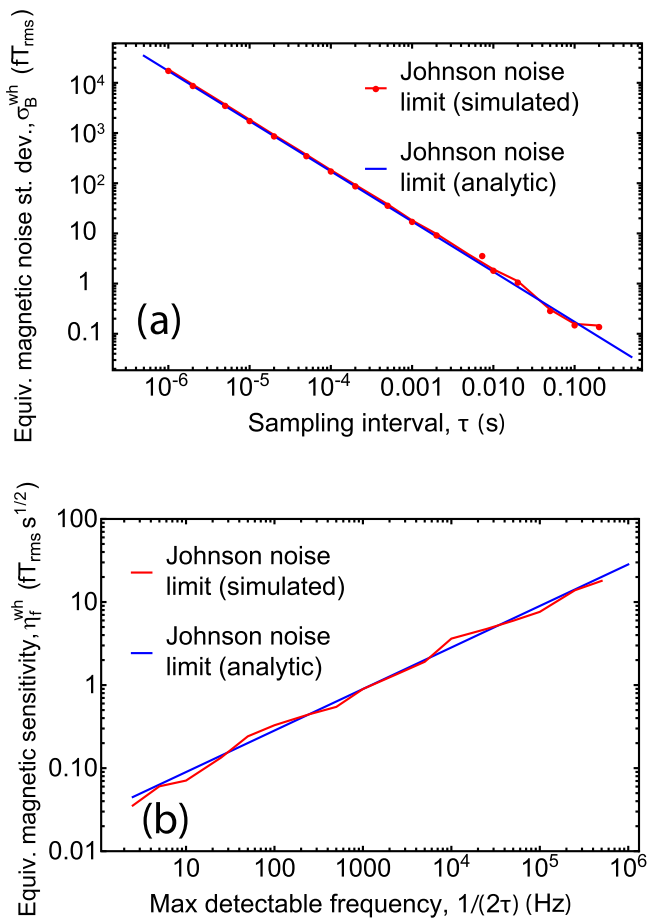

FIG. 11. Oscillator Johnson phase-noise limit: cw. (a) Equivalent magnetic noise standard deviation due to room-temperature oscillator Johnson phase noise, σ_B^{wh}, versus sampling interval τ. (b) Equivalent magnetic sensitivity due to room-temperature oscillator Johnson phase noise, η_f^{wh}, as a function of magnetometer bandwidth, $1/(2\tau)$.

bandwidth magnetometer, the equivalent magnetic sensitivity is 1.4 $\mathrm{pT_{rms}\,s^{1/2}}$ for G1 and 0.3 $\mathrm{pT_{rms}\,s^{1/2}}$ for G2.

1. Oscillator Johnson phase-noise limit: Continuous wave

For white phase noise, such as that due to oscillator Johnson phase noise, $S_\phi(f) = S_\phi^{wh}$. The integral in Eq. (I2) can be done explicitly, resulting in a frequency-noise standard deviation

$$\sigma_f(\tau) \approx \frac{\sqrt{2S_\phi^{wh} f_c}}{2\pi\tau}. \tag{I5}$$

Inserting Eq. (I2) into Eq. (I3), the equivalent magnetic-noise standard deviation for a single sampling interval τ is given by

$$\sigma_B^{wh}(\tau) \approx \frac{\sqrt{2S_\phi^{wh} f_c}}{2\pi\gamma_{\mathrm{nv}}\tau}. \tag{I6}$$

Figure 11(a) shows plots of $\sigma_B(\tau)$, using either Eq. (I6) ("analytic") or by directly propagating numerically simulated phase noise with Eq. (I1) ("simulated"). In either case, we assume a room-temperature Johnson-noise-limited, 0 dBm oscillator phase noise, $S_\phi^{wh} = 2 \times 10^{-17.7}$ rad^2/Hz, and $f_c = 1$ MHz. The linear dependence of $\sigma_B(\tau)$ on τ for white phase noise implies that longer sampling intervals offer superior sensitivity. However, the maximum detectable magnetic-field frequency is $1/(2\tau)$, so long sampling intervals restrict the bandwidth of the magnetometer.

Incorporating Eq. (I6) in Eq. (I4), the equivalent magnetic sensitivity in the room-temperature, oscillator-Johnson-phase-noise limit is

$$\eta_f^{wh}(\tau) = \sigma_B(\tau)\sqrt{\tau} \approx \frac{1}{\pi\gamma_{\mathrm{nv}}}\sqrt{\frac{S_\phi^{wh} f_c}{2\tau}}. \tag{I7}$$

Figure 11(b) shows a plot of $\eta_f^{wh}(\tau)$ as a function of the maximum detectable frequency, $1/(2\tau)$. The numerically simulated values are formed by taking 1-s intervals of simulated noise, resampling by averaging over consecutive intervals of $\tau \leqslant 0.5$ s, and taking the maximum frequency component. The noise is then the standard deviation of 50 of these 1-s intervals. At low frequencies, the equivalent magnetic sensitivity due to Johnson phase noise is negligible, but for

frequencies $\gtrsim$100 kHz the noise is $\gtrsim$10 $\mathrm{fT_{rms}\,s^{1/2}}$, which is above fundamental limits [25].

2. Lock-in measurements

We anticipate that the above analysis should also hold for experiments using a lock-in amplifier. Our primary assumption is a linear relationship between the fluorescence signal and the MW carrier frequency detuning, Eq. (I1), which should also hold true for a lock-in measurement. One caveat is that, when filtering is done in a lock-in experiment, there is a minimum effective sampling interval τ_{min} given by $1/(2\tau_{\mathrm{min}}) \approx f_{\mathrm{enbw}} \lesssim f_{\mathrm{mod}}$, where f_{enbw} is the effective noise bandwidth and f_{mod} is the modulation frequency of the lock-in experiment.

3. Suppressing the impact of microwave phase noise in a continuous-wave experiment

As mentioned in Sec. VI of the main text, we expect that the dual-resonance scheme, first described in Ref. [43] (see also Ref. [14]), should be effective at suppressing frequency noise in cw optically detected magnetic resonance measurements. Consider the MW mixer schematic presented in Appendix K with the "V" level diagram. The MW Rabi frequency and optical pump rate are assumed to be weak enough, relative to the spin-resonance linewidth, to avoid populating coherent dark states [57]. If the output frequencies of the mixer are tuned to drive opposing sides of the $f_{\pm}$ resonances, the impact of the frequency noise of the main oscillator, with carrier frequency $D = 2.87$ GHz, on the NV signal should cancel. The remaining frequency noise from the local oscillator does not cancel, but its carrier frequency is much smaller, $\approx\gamma_{\mathrm{nv}}B_0 \ll D$. If the frequency noise scales linearly with carrier frequency, the result should be an improvement in the frequency-noise-limited sensitivity by a factor $\approx D/(\gamma_{\mathrm{nv}}B_0)$.

APPENDIX J: TIME-DOMAIN MODEL OF PHASE-NOISE-LIMITED SENSITIVITY

In this section, formulas are derived for the equivalent magnetic-field noise resulting from MW pulse sequences with white and random-walk phase noise. We focus on pulse sequences using a series of N resonant MW π pulses with uniform spacing, such as CPMG and XY pulse sequences. However, a similar analysis can be applied to a wide variety of pulse sequences. The MW pulses are in resonance with only one of the NV $f_{\pm}$ spin transitions such that the spin dynamics can be described by an effective spin-1/2 system. We assume all pulses are instantaneous, resonant pulses and there are no additional time-varying fields.

Figure 12 shows an example CPMG pulse sequence along with the associated phase of the NV spin state on the equator of the Bloch sphere. NV centers are initially polarized along the z axis of the Bloch sphere. A MW $\pi/2$ pulse along the x axis rotates the NV electron spin to lie along the y axis. The phase of this first $\pi/2$ pulse defines the coordinates of the Bloch sphere for all subsequent pulses. At this point, the NV phase is defined to be $\phi_0 = 0$ with respect to the y axis, and subsequent errors in the microwave phase are defined relative to this axis. Next, N resonant π pulses, spaced a time 2τ apart,

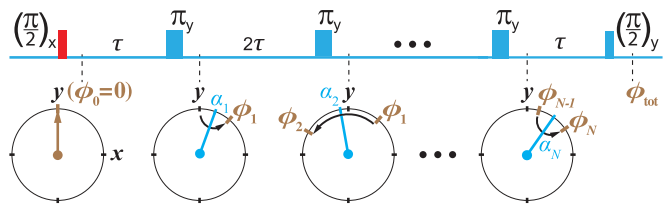

FIG. 12. Propagation of spin state rotations due to MW phase errors. A graphical representation of the model for a spin-1/2 qubit under a multipulse (N π pulses) sequence in the absence of time-varying magnetic fields. All pulses are assumed to be ideal, resonant pulses of negligible length. The first MW $\pi/2$ pulse rotates the NV spin to lie along the y axis (Bloch vector angle: $\phi_0 = 0$), where the x and y axes are defined by the MW carrier's phase during this pulse. For subsequent MW π pulses, the spin state is reflected about a rotated axis of angle α_i with respect to the y axis, where α_i is the MW phase error. The spin state's phase after the ith pulse is $\phi_i = 2\alpha_i - \phi_{i-1}$, see Eq. (J1).

are applied. While the π pulses are applied approximately along the y axis, each pulse has a small phase error, α_i, with respect to the y axis. The effect of each π pulse is to apply a π rotation of the NV spin state about the axis defined by the MW phase error α_i. In the absence of external magnetic signals, the phase of the NV spin state ϕ_i after the ith π pulse can be written as

$$\phi_i = 2(\alpha_i - \phi_{i-1}) + \phi_{i-1} = 2\alpha_i - \phi_{i-1}, \tag{J1}$$

where ϕ_{i-1} is the phase of the spin state prior to the pulse. The total phase accumulated by the NV spin state, ϕ_{tot}, at the end of the pulse sequence (Fig. 12) is given by

$$\phi_{\mathrm{tot}} = -\alpha_f + \phi_N = -\alpha_f + \sum_{i=1}^{N} (-1)^{N-i} 2\alpha_i, \tag{J2}$$

where α_f is the MW phase error of the final $\pi/2$ pulse.

1. White microwave phase noise

Equation (J2) shows how MW phase errors lead to an error in the NV spin state's phase. We now consider the case when the MW phase errors, α_i, are normally distributed, as is the case for white noise. White phase noise is especially prominent at higher MW carrier offset, as seen in Fig. 2(d). A fundamental constraint on the phase-noise performance of MW generators is due to Johnson noise in the MW oscillator which has a white-noise character.

The variance of the NV total phase error, $\mathrm{Var}(\phi_{\mathrm{tot}})$, is the sum of the variances of each term on the right side of Eq. (J2):

$$\mathrm{Var}(\phi_{\mathrm{tot}}) = \mathrm{Var}(\alpha_f) + \sum_{i=1}^{N} \mathrm{Var}(2\alpha_i). \tag{J3}$$

In the case of white noise, the variance of the MW phase errors, $\mathrm{Var}(\alpha_{f,i})$, is independent of the temporal spacing between the pulses and only depends on the standard deviation σ_{wh} of the MW phase error in each pulse. In other words,

$$\mathrm{Var}(\alpha_i) = \mathrm{Var}(\alpha_f) = \sigma_{\mathrm{wh}}^2. \tag{J4}$$

Inserting Eq. (J4) in Eq. (J3), the total variance becomes

$$\mathrm{Var}(\phi_{\mathrm{tot}}) = 4\sigma_{\mathrm{wh}}^2(N + 1/4), \tag{J5}$$

and the standard deviation of the total phase error of the NV spin state, σ_ϕ, is

$$\sigma_\phi = \sqrt{\mathrm{Var}(\phi_{\mathrm{tot}})} = 2\sigma_{\mathrm{wh}}\sqrt{N + 1/4}. \tag{J6}$$

Inserting Eq. (J6) in Eq. (2), using $\tau_{\mathrm{tot}} = 2N\tau$ and assuming $N \gg 1$, the white-phase-noise-limited equivalent magnetic sensitivity can be written as

$$\eta_{\mathrm{wh}} \approx \frac{\sigma_{\mathrm{wh}}\sqrt{N(1 + t_{\mathrm{dead}}/\tau_{\mathrm{tot}})}}{2\gamma_{\mathrm{nv}}\sqrt{\tau_{\mathrm{tot}}}} = \frac{\sigma_{\mathrm{wh}}}{\gamma_{\mathrm{nv}}}\sqrt{\frac{f_{\mathrm{xy8}}}{2\delta}}, \tag{J7}$$

where $f_{\mathrm{xy8}} = 1/(4\tau)$ is the central frequency of the pulse sequence's filter function (we use the subscript xy8 here, but the same definition applies for CPMG sequences) and $\delta = \tau_{\mathrm{tot}}/(\tau_{\mathrm{tot}} + t_{\mathrm{dead}})$ is the duty cycle of the NV readouts.

2. Random-walk microwave phase noise

If the MW phase errors (α_i) are dominated by random-walk noise, the relative MW phase error accumulates. The phase error of the ith pulse is given by

$$\alpha_i = \alpha_{i-1} + \Delta_i, \tag{J8}$$

where Δ_i is the MW phase displacement accrued during the time interval between pulses $i - 1$ and i. Similar to the white-noise case, we suppose that the first $\pi/2$ pulse in Fig. 12 is perfect, $\alpha_0 = 0$. Inserting Eq. (J8) into Eq. (J2), the total phase error of the NV spin state for even and odd numbers of the applied π pulses can be written as

$$\phi_{\mathrm{tot}} = \begin{cases} -\alpha_f + \sum_{i=1}^{N/2} 2\Delta_{2i}, & N \text{ even} \\ -\alpha_f + \sum_{i=1}^{(N+1)/2} 2\Delta_{2i-1}, & N \text{ odd.} \end{cases} \tag{J9}$$

This can be further reduced to

$$\phi_{\mathrm{tot}} = -\Delta_f + \sum_{i=1}^{N} (-1)^{N-i}\Delta_i. \tag{J10}$$

In Eq. (J10), Δ_1 and Δ_f are special cases where the phase displacement is accrued over an interval τ (all other Δ_i accrue over an interval 2τ), and their variances can be assumed to be the same. Specifically,

$$\mathrm{Var}(\Delta_1) = \mathrm{Var}(\Delta_f) = \sigma_{\mathrm{rw}}^2 R_{\mathrm{samp}}\tau. \tag{J11}$$

Here we have modeled the random-walk MW phase evolution as a set of discrete jumps in MW phase, where σ_{rw} is the standard deviation of the phase jumps and $R_{\mathrm{samp}} \gtrsim 1/(2\tau)$ is the jump sampling rate.

Because the π pulses are evenly spaced in time by 2τ (see Fig. 12), the variance of the random-walk MW phase displacement of all other Δ_i ($i = 2, 3, \ldots N$) is two times larger than those in Eq. (J11):

$$\mathrm{Var}(\Delta_2) = \mathrm{Var}(\Delta_3) = \cdots = \mathrm{Var}(\Delta_N) = \sigma_{\mathrm{rw}}^2 R_{\mathrm{samp}} 2\tau. \tag{J12}$$

Using Eqs. (J11) and Eq. (J12) with Eq. (J10), the variance of the phase of the NV spin state due to MW random-walk phase noise becomes

$$\mathrm{Var}(\phi_{\mathrm{tot}}) = \sigma_{\mathrm{rw}}^2 2N\tau R_{\mathrm{samp}}, \tag{J13}$$

and its standard deviation is

$$\sigma_\phi = \sigma_{\mathrm{rw}}\sqrt{2N\tau R_{\mathrm{samp}}} = \sigma_{\mathrm{rw}}\sqrt{\tau_{\mathrm{tot}} R_{\mathrm{samp}}}. \tag{J14}$$

Inserting Eq. (J14) into Eq. (2) of the main text, the random-walk-phase-noise-limited equivalent magnetic sensitivity is given by

$$\eta_{\mathrm{rw}} \approx \frac{\sigma_{\mathrm{rw}}}{4\gamma_{\mathrm{nv}}}\sqrt{\frac{R_{\mathrm{samp}}}{\delta}}. \tag{J15}$$

3. Extension to XY8-N_{r} pulse sequences

In the case of the XY8-N_r pulse sequence used in this study, the total number of the applied π pulses is $N = 8N_{\mathrm{r}}$. The total NV spin-state phase accumulation due to MW phase errors is given by

$$\phi_{\mathrm{tot}} = -\alpha_f + \sum_{i=1}^{N} (-1)^i 2\alpha_i, \tag{J16}$$

which is identical to the expression for N even derived for the CPMG sequence, Eq. (J1). Here the MW phase errors α_i are with respect to the desired phase. For XY8 sequences, the desired phase alternates between 0° and 90°, but it does so in such a way that there is no final phase accumulation in the absence of spin precession or phase errors. Thus the expressions derived above for σ_ϕ [Eqs. (J6) and (J14)], and η_ϕ [Eqs. (J7) and (J15)] are also valid for the XY8-N_{r} sequence.

APPENDIX K: PHASE NOISE IN DOUBLE-QUANTUM MEASUREMENT

In the case of double-quantum measurements, two microwave fields with the same amplitude are simultaneously applied at both $f_\pm$ transition frequencies. Here we assume the bias magnetic field is small, $B_0 \ll D/\gamma_{\mathrm{nv}}$.

Figure 13 depicts a double-quantum Ramsey pulse sequence. The first two-tone resonant MW pulse transfers the NV population from $|m_{\mathrm{s}} = 0\rangle$ state to the "bright" coherent superposition state $|\Psi_{\mathrm{B}}\rangle$ given by [58,59]

$$|\Psi_{\mathrm{B}}\rangle = \frac{1}{\sqrt{2}} e^{-i\frac{\alpha_1+\alpha_2}{2}} [e^{i\frac{\Delta\alpha}{2}}|-1\rangle + e^{-i\frac{\Delta\alpha}{2}}|+1\rangle]. \tag{K1}$$

In Eq. (K1), α_1 and α_2 are the phases of the two resonant MW tones, and $\Delta\alpha = \alpha_2 - \alpha_1$.

If the second two-tone pulse, with MW phases α_1' and α_2', is applied in a time interval τ (Fig. 13), it will project the NV

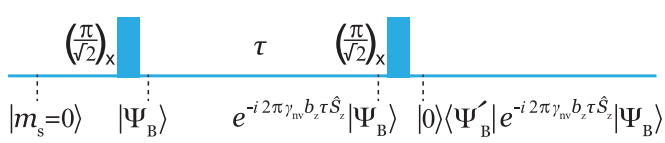

FIG. 13. Double-quantum Ramsey sequence. Microwave pulses are formed from two tones resonant with both $f_\pm$ transitions. The first MW pulse transfers the NV population from $|m_{\mathrm{s}} = 0\rangle$ state to the "bright" coherent superposition state $|\Psi_{\mathrm{B}}\rangle$, Eq. (K1). The state then freely evolves under a small magnetic field b_z for a time τ. A second MW pulse transfers the bright state $|B'\rangle$ back to $|m_{\mathrm{s}} = 0\rangle$ state. Note that $|B\rangle \neq |B'\rangle$ if the relative phase of the MW tones changes, see Eqs. (K2) and (K1).

population back into the $|0\rangle$ state with a probability given by

$$|\langle\Psi_B'|^{-i2\pi\gamma_{nv}b_z\tau\hat{S}_z}|\Psi_B\rangle|^2 = \cos^2\left(\frac{\Delta\alpha'-\Delta\alpha}{2}-2\pi\gamma_{nv}b_z\tau\right), \tag{K2}$$

where $\Delta\alpha' = \alpha_2' - \alpha_1'$. In Eq. (K2), $b_z \ll \Omega_R/\gamma_{nv}$ is a small magnetic field to be detected that has a negligible effect on the NV coherent bright states, Eq. (K1). If this magnetic field is very small, $2\pi\gamma_{nv}b_z\tau \ll 1$, then $\Delta\alpha'$ can be offset by 90° to ensure the probability in Eq. (K2) depends linearly on b_z.

Figure 14(a) shows a method to generate the two MW tones that minimizes the impact of phase noise [14,42]. Figure 14(b) shows the associated NV level diagrams in the low-field regime. By mixing a lower-frequency signal generator as the local oscillator [$f_{LO} = (f_+ - f_-)/2 \approx \gamma_{nv}B_0$] and a higher-frequency signal generator as the carrier [$f_{car} = (f_+ + f_-)/2 \approx D$], both the NV spin transitions with frequencies $f_\pm$ can be driven simultaneously. Moreover, the phase noise of the higher-frequency f_{car} tone has no impact on the magnetometer signal [see Eq. (K2) and Appendix G], leaving only the phase noise due to f_{LO}. If the phase noise scales linearly with MW carrier frequency, this scheme can suppress the impact of MW phase noise by a factor $\approx D/\gamma_{nv}B_0$ in low fields ($\gamma_{nv}B_0 \ll D$).

This measurement protocol is analogous to how high-precision atomic spectroscopy of three-level atoms is conducted in the optical regime with noisy lasers [60]. At very low applied field $B_0 \lesssim \Omega_R/\gamma_{nv}$, where Ω_R is the MW Rabi frequency, a single MW tone can be used with a double-quantum pulse sequence [44], potentially eliminating the first-order impact of MW phase noise entirely. At high magnetic field, $B_0 > D/\gamma_{nv}$, this method does not appear to lead to a phase-noise cancellation, and alternative methods should be pursued.

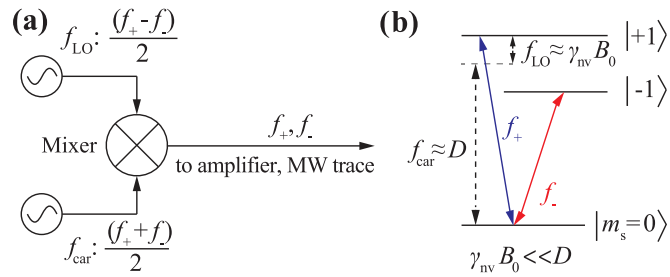

FIG. 14. Double-quantum and dual-resonance schemes. (a) Electronic schematic for generating dual-tone MW frequencies in resonant with NV spin transitions $f_\pm$ [14,42]. f_{LO}: frequency of local oscillator, f_{car}: frequency of carrier generator. (b) Ground-state NV spin level diagram for the "V" scheme ($B_0 \ll D/\gamma_{nv}$).

[1] D. R. Glenn, K. Lee, H. Park, R. Weissleder, A. Yacoby, M. D. Lukin, H. Lee, R. L. Walsworth, and C. B. Connolly, Single-cell magnetic imaging using a quantum diamond microscope, Nat. Methods **12**, 736 (2015).

[2] I. Fescenko, A. Laraoui, J. Smits, N. Mosavian, P. Kehayias, J. Seto, L. Bougas, A. Jarmola, and V. M. Acosta, Diamond magnetic microscopy of malarial hemozoin nanocrystals, Phys. Rev. Appl. **11**, 034029 (2019).

[3] P. Wang, S. Chen, M. Guo, S. Peng, M. Wang, M. Chen, W. Ma, R. Zhang, J. Su, X. Rong, F. Shi, T. Xu, and J. Du, Nanoscale magnetic imaging of ferritins in a single cell, Sci. Adv. **5**, eaau8038 (2019).

[4] J. M. McCoey, M. Matsuoka, R. W. de Gille, L. T. Hall, J. A. Shaw, J.-P. Tetienne, D. Kisailus, L. C. L. Hollenberg, and D. A. Simpson, Diamond-based magnetic microscopy: Quantum magnetic imaging of iron biomineralization in teeth of the chiton acanthopleura hirtosa, Small Methods **4**, 2070010 (2020).

[5] J. L. Webb, L. Troise, N. W. Hansen, C. Olsson, A. M. Wojciechowski, J. Achard, O. Brinza, R. Staacke, M. Kieschnick, J. Meijer, A. Thielscher, J.-F. Perrier, K. Berg-Sorensen, A. Huck, and U. L. Andersen, Detection of biological signals from a live mammalian muscle using an early stage diamond quantum sensor, Sci. Rep. **11**, 2412 (2021).

[6] Z. Kazi, I. M. Shelby, R. Nirodi, J. Turnbull, H. Watanabe, K. M. Itoh, P. A. Wiggins, and K.-M. C. Fu, Direct measure of DNA bending by quantum magnetic imaging of a nano-mechanical torque-balance, arXiv:2402.17964.

[7] N. Aslam, M. Pfender, P. Neumann, R. Reuter, A. Zappe, F. F. de Oliveira, A. Denisenko, H. Sumiya, S. Onoda, J. Isoya, and J. Wrachtrup, Nanoscale nuclear magnetic resonance with chemical resolution, Science **357**, 67 (2017).

[8] D. R. Glenn, D. B. Bucher, J. Lee, M. D. Lukin, H. Park, and R. L. Walsworth, High-resolution magnetic resonance spectroscopy using a solid-state spin sensor, Nature (London) **555**, 351 (2018).

[9] J. Smits, J. T. Damron, P. Kehayias, A. F. McDowell, N. Mosavian, I. Fescenko, N. Ristoff, A. Laraoui, A. Jarmola, and V. M. Acosta, Two-dimensional nuclear magnetic resonance spectroscopy with a microfluidic diamond quantum sensor, Sci. Adv. **5**, eaaw7895 (2019).

[10] F. Bruckmaier, R. D. Allert, N. R. Neuling, P. Amrein, S. Littin, K. D. Briegel, P. Schätzle, P. Knittel, M. Zaitsev, and D. B. Bucher, Imaging local diffusion in microstructures using NV-based pulsed field gradient NMR, Sci. Adv. **9**, eadh3484 (2023).

[11] J.-C. Jaskula, K. Saha, A. Ajoy, D. J. Twitchen, M. Markham, and P. Cappellaro, Cross-sensor feedback stabilization of an emulated quantum spin gyroscope, Phys. Rev. Appl. **11**, 054010 (2019).

[12] V. V. Soshenko, S. V. Bolshedvorskii, O. Rubinas, V. N. Sorokin, A. N. Smolyaninov, V. V. Vorobyov, and A. V. Akimov, Nuclear spin gyroscope based on the nitrogen vacancy center in diamond, Phys. Rev. Lett. **126**, 197702 (2021).

[13] A. Jarmola, S. Lourette, V. M. Acosta, A. G. Birdwell, P. Blümler, D. Budker, T. Ivanov, and V. S. Malinovsky, Demonstration of diamond nuclear spin gyroscope, Sci. Adv. **7**, eabl3840 (2021).

[14] I. Fescenko, A. Jarmola, I. Savukov, P. Kehayias, J. Smits, J. Damron, N. Ristoff, N. Mosavian, and V. M. Acosta, Diamond magnetometer enhanced by ferrite flux concentrators, Phys. Rev. Res. **2**, 023394 (2020).

[15] Y. Silani, J. Smits, I. Fescenko, M. W. Malone, A. F. McDowell, A. Jarmola, P. Kehayias, B. A. Richards, N. Mosavian, N.

Ristoff, and V. M. Acosta, Nuclear quadrupole resonance spectroscopy with a femtotesla diamond magnetometer, Sci. Adv. **9**, eadh3189 (2023).

[16] X. Gao, C. Yu, S. Zhang, H. Lin, J. Guo, M. Ma, Z. Feng, and F.-W. Sun, High sensitivity of diamond nitrogen-vacancy magnetometer with magnetic-flux concentrators via enhanced fluorescence collection, Diamond Relat. Mater. **139**, 110348 (2023).

[17] J. F. Barry, M. H. Steinecker, S. T. Alsid, J. Majumder, L. M. Pham, M. F. O'Keefe, and D. A. Braje, Sensitive AC and DC magnetometry with nitrogen-vacancy center ensembles in diamond, arXiv:2305.06269.

[18] X. Rong, M. Wang, J. Geng, X. Qin, M. Guo, M. Jiao, Y. Xie, P. Wang, P. Huang, F. Shi, Y.-F. Cai, C. Zou, and J. Du, Searching for an exotic spin-dependent interaction with a single electron-spin quantum sensor, Nat. Commun. **9**, 739 (2018).

[19] P.-H. Chu, N. Ristoff, J. Smits, N. Jackson, Y. J. Kim, I. Savukov, and V. M. Acosta, Proposal for the search for new spin interactions at the micrometer scale using diamond quantum sensors, Phys. Rev. Res. **4**, 023162 (2022).

[20] L. Wu, S. Lin, X. Kong, M. Wang, J. Zhou, C.-K. Duan, P. Huang, L. Zhang, and J. Du, A spin-mechanical quantum chip for exploring exotic interactions, Proc. Natl. Acad. Sci. USA **120**, e2302145120 (2023).

[21] F. Casola, T. van der Sar, and A. Yacoby, Probing condensed matter physics with magnetometry based on nitrogen-vacancy centres in diamond, Nat. Rev. Mater. **3**, 17088 (2018).

[22] T. Zhang, G. Pramanik, K. Zhang, M. Gulka, L. Wang, J. Jing, F. Xu, Z. Li, Q. Wei, P. Cigler, and Z. Chu, Toward quantitative bio-sensing with nitrogen–vacancy center in diamond, ACS Sens. **6**, 2077 (2021).

[23] E. Marchiori, L. Ceccarelli, N. Rossi, L. Lorenzelli, C. L. Degen, and M. Poggio, Nanoscale magnetic field imaging for 2D materials, Nat. Rev. Phy **4**, 49 (2022).

[24] N. Aslam, H. Zhou, E. K. Urbach, M. J. Turner, R. L. Walsworth, M. D. Lukin, and H. Park, Quantum sensors for biomedical applications, Nat. Rev. Phy **5**, 157 (2023).

[25] 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).

[26] 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).

[27] H. Ball, W. D. Oliver, and M. J. Biercuk, The role of master clock stability in quantum information processing, npj Quantum Inf. **2**, 16033 (2016).

[28] E. Rubiola, *Phase Noise and Frequency Stability in Oscillators* (Cambridge University Press, 2008).

[29] J. F. Barry, M. J. Turner, J. M. Schloss, D. R. Glenn, Y. Song, M. D. Lukin, H. Park, and R. L. Walsworth, Optical magnetic detection of single-neuron action potentials using quantum defects in diamond, Proc. Natl. Acad. Sci. USA **113**, 14133 (2016).

[30] M. I. Ibrahim, C. Foy, D. R. Englund, and R. Han, High-scalability CMOS quantum magnetometer with spin-state excitation and detection of diamond color centers, IEEE J. Solid-State Circuits **56**, 1001 (2021).

[31] F. A. Hahl, L. Lindner, X. Vidal, T. Luo, T. Ohshima, S. Onoda, S. Ishii, A. M. Zaitsev, M. Capelli, B. C. Gibson, A. D. Greentree, and J. Jeske, Magnetic-field-dependent stimulated emission from nitrogen-vacancy centers in diamond, Sci. Adv. **8**, eabn7192 (2022).

[32] C. Zhang, J. Zhang, M. Widmann, M. Benke, M. Kübler, D. Dasari, T. Klotz, L. Gizzi, O. Röhrle, P. Brenner, and J. Wrachtrup, Optimizing NV magnetometry for magnetoneurography and magnetomyography applications, Front. Neurosci. **16**, (2023).

[33] N. Sekiguchi, M. Fushimi, A. Yoshimura, C. Shinei, M. Miyakawa, T. Taniguchi, T. Teraji, H. Abe, S. Onoda, T. Ohshima, M. Hatano, M. Sekino, and T. Iwasaki, Diamond quantum magnetometer with dc sensitivity of < 10 pT $\mathrm{Hz}^{-1/2}$ toward measurement of biomagnetic field, Phys. Rev. Appl. **21**, 064010 (2024).

[34] SRS SG386 data sheet (last accessed on 15 May, 2024), https://www.thinksrs.com/downloads/pdfs/manuals/SG380m.pdf.

[35] Rohde & Schwarz SMU200A data sheet (last accessed on 15 May, 2024), https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_common_library/dl_brochures_and_datasheets/pdf_1/SMU_dat-sw-en.pdf.

[36] S. M. Graham, A. J. Newman, C. J. Stephen, A. M. Edmonds, D. J. Twitchen, M. L. Markham, and G. W. Morley, On the road with a diamond magnetometer, arXiv:2401.16090.

[37] M. Biercuk, A. Doherty, and H. Uys, Dynamical decoupling sequence construction as a filter-design problem, J. Phys. B: At., Mol. Opt. Phys. **44**, 154002 (2011).

[38] G. A. Álvarez and D. Suter, Measuring the spectrum of colored noise by dynamical decoupling, Phys. Rev. Lett. **107**, 230501 (2011).

[39] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Noise spectroscopy through dynamical decoupling with a superconducting flux qubit, Nat. Phys. **7**, 565 (2011).

[40] N. Bar-Gill, L. M. Pham, C. Belthangady, D. Le Sage, P. Cappellaro, J. R. Maze, M. D. Lukin, A. Yacoby, and R. Walsworth, Suppression of spin-bath dynamics for improved coherence of multi-spin-qubit systems, Nat. Commun. **3**, 858 (2012).

[41] A. Hati, D. Howe, F. Walls, and D. Walker, Noise figure vs. PM noise measurements: A study at microwave frequencies, in *Proceedings of the 2003 IEEE International Frequency Control Symposium and PDA Exhibition Jointly with the 17th European Frequency and Time Forum* (IEEE, Piscataway, NJ, 2003), pp. 516–520.

[42] H. J. Mamin, M. H. Sherwood, M. Kim, C. T. Rettner, K. Ohno, D. D. Awschalom, and D. Rugar, Multipulse double-quantum magnetometry with near-surface nitrogen-vacancy centers, Phys. Rev. Lett. **113**, 030803 (2014).

[43] A. M. Wojciechowski, M. Karadas, C. Osterkamp, S. Jankuhn, J. Meijer, F. Jelezko, A. Huck, and U. L. Andersen, Precision temperature sensing in the presence of magnetic field noise and vice-versa using nitrogen-vacancy centers in diamond, Appl. Phys. Lett. **113**, 013502 (2018).

[44] K. Fang, V. M. Acosta, C. Santori, Z. Huang, K. M. Itoh, H. Watanabe, S. Shikata, and R. G. Beausoleil, High-sensitivity magnetometry based on quantum beats in diamond nitrogen-vacancy centers, Phys. Rev. Lett. **110**, 130802 (2013).

[45] D. Chaudy, O. Llopis, B. Marcilhac, Y. Lemaître, O. d. Kelly, J.-M. Hode, and J.-M. Lesage, A low phase noise all cryogenic microwave oscillator based on a superconductor resonator, IEEE Trans. Ultrson. Ferroelectr. Freq. Control **67**, 2750 (2020).

[46] E. N. Ivanov and M. E. Tobar, Low phase-noise sapphire crystal microwave oscillators: Current status, IEEE Trans. Ultrson. Ferroelectr. Freq. Control **56**, 263 (2009).

[47] A. Hati, C. W. Nelson, B. Riddle, and D. A. Howe, Pm noise of a 40 GHz air-dielectric cavity oscillator, arXiv:1404.4828.

[48] M. Goryachev, W. G. Farr, D. L. Creedon, Y. Fan, M. Kostylev, and M. E. Tobar, High-cooperativity cavity QED with magnons at microwave frequencies, Phys. Rev. Appl. **2**, 054002 (2014).

[49] J. F. Barry, R. A. Irion, M. H. Steinecker, D. K. Freeman, J. J. Kedziora, R. G. Wilcox, and D. A. Braje, Ferrimagnetic oscillator magnetometer, Phys. Rev. Appl. **19**, 044044 (2023).

[50] X. S. Yao and L. Maleki, Optoelectronic microwave oscillator, J. Opt. Soc. Am. B **13**, 1725 (1996).

[51] J. Li and K. Vahala, Small-sized, ultra-low phase noise photonic microwave oscillators at X-Ka bands, Optica **10**, 33 (2023).

[52] I. Kudelin, W. Groman, Q.-X. Ji, J. Guo, M. L. Kelleher, D. Lee, T. Nakamura, C. A. McLemore, P. Shirmohammadi, S. Hanifi, H. Cheng, N. Jin, L. Wu, S. Halladay, Y. Luo, Z. Dai, W. Jin, J. Bai, Y. Liu, W. Zhang *et al.*, Photonic chip-based low-noise microwave oscillator, Nature (London) **627**, 534 (2024).

[53] J. Shao, Y. Luo, J. Chen, H. Huang, G.-S. Liu, L. Chen, Z. Chen, and Y. Chen, High-sensitivity optical-fiber magnetic sensor based on diamond and magnetic flux concentrators, Opt. Express **31**, 14685 (2023).

[54] P. Kehayias, A. Jarmola, N. Mosavian, I. Fescenko, F. M. Benito, A. Laraoui, J. Smits, L. Bougas, D. Budker, A. Neumann, S. R. J. Brueck, and V. M. Acosta, Solution nuclear magnetic resonance spectroscopy on a nanostructured diamond chip, Nat. Commun. **8**, 188 (2017).

[55] R. Lazda, L. Busaite, A. Berzins, J. Smits, F. Gahbauer, M. Auzinsh, D. Budker, and R. Ferber, Cross-relaxation studies with optically detected magnetic resonances in nitrogen-vacancy centers in diamond in external magnetic field, Phys. Rev. B **103**, 134104 (2021).

[56] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. G. 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).

[57] J. H. Shim, S.-J. Lee, S. Ghimire, J. I. Hwang, K.-G. Lee, K. Kim, M. J. Turner, C. A. Hart, R. L. Walsworth, and S. Oh, Multiplexed sensing of magnetic field and temperature in real time using a nitrogen-vacancy ensemble in diamond, Phys. Rev. Appl. **17**, 014009 (2022).

[58] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, 1997).

[59] V. Shah and J. Kitching, Advances in coherent population trapping for atomic clocks, in *Advances in Atomic, Molecular, and Optical Physics* (Elsevier, 2010), Vol. 59, pp. 21–74.

[60] J. E. Thomas, P. R. Hemmer, S. Ezekiel, C. C. Leiby, R. H. Picard, and C. R. Willis, Observation of Ramsey fringes using a stimulated, resonance Raman transition in a sodium atomic beam, Phys. Rev. Lett. **48**, 867 (1982).