Editors' Suggestion

Sequential Bayesian Experiment Design for Optically Detected Magnetic Resonance of Nitrogen-Vacancy Centers

Sergey Dushenko,[1,2,*] Kapildeb Ambal,[1,2,3] and Robert D. McMichael[1]

[1] *Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*

[2] *Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA*

[3] *Department of Physics, Wichita State University, Wichita, Kansas 67260, USA*

(Received 31 July 2020; revised 1 October 2020; accepted 9 October 2020; published 16 November 2020)

In magnetometry using optically detected magnetic resonance of nitrogen vacancy (N-V^-) centers, we demonstrate speedup of more than 1 order of magnitude with a sequential Bayesian experiment design as compared with conventional frequency-swept measurements. The N-V^- center is an excellent platform for magnetometry, with potential spatial resolution down to a few nanometers and demonstrated single-defect sensitivity down to nanoteslas per square root hertz. The N-V^- center is a quantum defect with spin $S = 1$ and coherence time up to several milliseconds at room temperature. Zeeman splitting of the N-V^- energy levels allows detection of the magnetic field via photoluminescence. We compare conventional N-V^--center photoluminescence measurements that use predetermined sweeps of the microwave frequency with measurements using a Bayesian-inference method. In sequential Bayesian experiment design, the settings with maximum utility are chosen for each measurement in real time on the basis of the accumulated experimental data. Using this method, we observe an order of magnitude decrease in the N-V^--center-magnetometry measurement time necessary to achieve a set precision.

DOI: 10.1103/PhysRevApplied.14.054036

I. INTRODUCTION

This study focuses on magnetometry using optically detected magnetic resonance of negatively charged nitrogen-vacancy (N-V^-) centers. The ability to optically prepare and manipulate spin states, along with a long spin lifetime and robustness with regard to the environment made N-V^- centers a promising platform for application in various areas. A few prominent examples include quantum computing [1], cryptography [2], and memory [3,4]; biocompatible markers [5] and drug delivery [6]; and mechanical [7], temperature [8,9], electric [10] and magnetic [11–13] sensors. The concept of N-V^--center magnetometry [14] was experimentally demonstrated in 2008 in independent studies by Balasubramanian *et al.* [11] and Maze *et al.* [12], followed by hundreds of other studies [15].

Magnetometry-based imaging using N-V^- centers promises several advantages over existing magnetic imaging and scanning techniques. The N-V^- center does not carry a significant magnetic moment, making N-V^--center magnetometry a noninvasive technique, unlike magnetic force microscopy, which can suffer from the interaction between the sample and the magnetic tip. Magneto-optic-Kerr-effect microscopy is limited by the optical resolution and is suitable mostly only for studying materials with a strong Kerr effect. In contrast, the spatial resolution of N-V^--center magnetometry is ultimately limited only by the distance between the N-V^- center and the sample, which can be less than 10 nm [16]. Superconducting-quantum-interference-device magnetometry provides unrivaled sensitivity but requires cryogenic temperatures, and has low spatial resolution, although attempts at miniaturizing the technology are in progress [17]. N-V^--center magnetometry can operate in a broad range of temperatures, including room temperature and above. These advantages make the N-V^- center an excellent platform for magnetometry [15,18,19], with potential spatial resolution down to a few nanometers and demonstrated sensitivity down to nanoteslas per square root hertz [20,21].

Recent research efforts have been directed at increasing the speed and precision of N-V^--center-magnetometry measurements. Some of these research efforts summon help from additional hardware to achieve the goal. By modulation of the microwave frequency that drives spin-state transitions of the N-V^- center and by demodulation of the photoluminescence signal with use of lock-in amplifiers, significant gains in the signal-to-noise ratio

*dushenko89@gmail.com

and measurement speed have been achieved [20,22–24]. However, such an approach generally requires a high photoluminescence signal by simultaneous measurement of multiple N-V^- centers, which sacrifices the spatial resolution. Another approach that uses specialized hardware is the use of a differential photon rate meter that can track the photoluminescence signal even at a low photon count rate, although it does not significantly increase the signal-to-noise ratio [25]. In addition to "hardware" approaches, sophisticated algorithms—"software" approaches—have also shown promise. Simulations have shown that neural networks improve N-V^--center readout fidelity [26]. Sequential Bayesian experiment design [27] is another promising machine-learning "software" approach. Theoretical studies have discussed how Bayesian methodology [28–31] can be used in determining the unknown parameters of a quantum system [32–36], and magnetometry in particular [37–40]. Encouragingly, in recent experimental studies, Bayesian methodology has proven to be advantageous in quantum Hamiltonian learning [41] and measurements of pulsed Ramsey magnetometry using N-V^- centers [42,43]. In this study, we show how combining sequential Bayesian experiment design with conventional optically-detected-magnetic-resonance N-V^--center magnetometry leads to better measurement strategies. In particular, we perform experiments that compare the use of a conventional swept-frequency N-V^--center-magnetometry protocol with measurements that incorporate sequential Bayesian experiment design.

II. BACKGROUND

Many of the useful properties of N-V^- centers hinge on the fact that their photoluminescence depends on their spin state. The N-V^- center is created when two adjacent carbon atoms in a diamond lattice are replaced by a vacancy and a negatively charged nitrogen atom, forming a spin $S = 1$ quantum defect [Fig. 1(a); see Sec. S.1 in Supplemental Material [44] for more details]. Photon absorption moves the N-V^- center from the ground state to the excited state, while preserving its spin projection m_S [Fig. 1(b)] [45,46]. Eventually, the center relaxes back to the ground state, but the relaxation process is spin dependent. An excited state with $m_S = 0$ mostly relaxes back to the ground state with $m_S = 0$ by emitting a red photon. In contrast, the excited state with $m_S = \pm 1$ can relax by two mechanisms: either back to the ground state with $m_S = \pm 1$, by emitting a red photon, or to any m_S through a dark state, without emitting a visible photon (the detailed energy-level structure of the N-V^- center can be found in Sec. S.1 in Supplemental Material [44]). Hence, photoluminescence of N-V^- centers under laser excitation is brighter if the center is initially in the $m_S = 0$ state and is dimmer if it is in the $m_S = \pm 1$ states. This phenomenon allows optical readout of the spin state by monitoring the photoluminescence rate. Additionally, the ground state with $m_S = 0$ of the N-V^- center can be prepared by continuous illumination that cycles N-V^- centers through ground-state–excited-state–ground-state transitions. Since the $m_S = \pm 1$ states can transition to the $m_S = 0$ state, but no reverse transition is available, eventually the center ends up in the $m_S = 0$ state with high probability. In all, the spin-dependent optical relaxation allows the spin state to be both initialized and read out.

The spin state of the N-V^- center can also be controlled with microwaves. When the microwave photon energy matches the energy difference between the ground levels with spin projection $m_S = 0$ and the $m_S = \pm 1$ spin states, transitions occur. The microwave energies at this resonance condition are given by

$$E_{\mathrm{MW}} = hf_{\mathrm{MW}} = hD_{\mathrm{GS}} + g\mu_B \Delta m_S B + m_I A_{\mathrm{GS}}^{\mathrm{HF}}, \quad (1)$$

where $h \approx 6.62 \times 10^{-34}$ J/Hz is the Planck constant, f_{MW} is the microwave frequency, $D_{\mathrm{GS}} \approx 2.87$ GHz is the zero-field splitting, $g \approx 2$ is the electron g factor inside the diamond lattice, $\mu_B \approx 9.27$ JT^{-1} is the Bohr magneton, Δm_S is the spin projection difference between the final and initial ground states, B is the applied magnetic field, m_I is the nuclear spin projection (preserved in the transition), and $A_{\mathrm{GS}}^{\mathrm{HF}}$ is the energy correction due to the hyperfine interaction of the ground-state levels with the ^{14}N nucleus (spin $I = 1$). Note that strain-induced splitting of the energy levels in diamond should also be considered when small magnetic fields below 1 mT are being measured.

Optically detected magnetic resonance [47,48] is observed as a reduction in photoluminescence. Constant illumination populates the $m_S = 0$ state, and dips in the photon count are observed when microwaves induce transitions to the $m_S = \pm 1$ states. One can extract the value of the external magnetic field B from the frequencies of the dips in the photoluminescence spectrum that correspond to the frequencies when the N-V^- center transitions to the $m_S = +1$ and $m_S = -1$ states [Fig. 1(c)]. This technique is the basis of N-V^--center magnetometry.

The resonance frequencies described in Eq. (1) yield a model for the normalized photon count signal ($y = \{\mu\}$) that is a combination of three Lorentzian curves, one for each of the ^{14}N nuclear I_z states in the hyperfine-interaction-split spectrum of the N-V^- center:

$$\mu = 1 - \frac{ak_{\mathrm{NP}}}{(f - f_B - \Delta f_{\mathrm{HF}})^2 + \Omega^2} - \frac{a}{(f - f_B)^2 + \Omega^2} - \frac{a/k_{\mathrm{NP}}}{(f - f_B + \Delta f_{\mathrm{HF}})^2 + \Omega^2}. \quad (2)$$

Here f_B is the center resonance frequency, which corresponds to the N-V^--center transition from the $\{m_S = 0, m_I = 0\}$ state to the $\{m_S = +1, m_I = 0\}$ state, $\Delta f_{\mathrm{HF}} =$

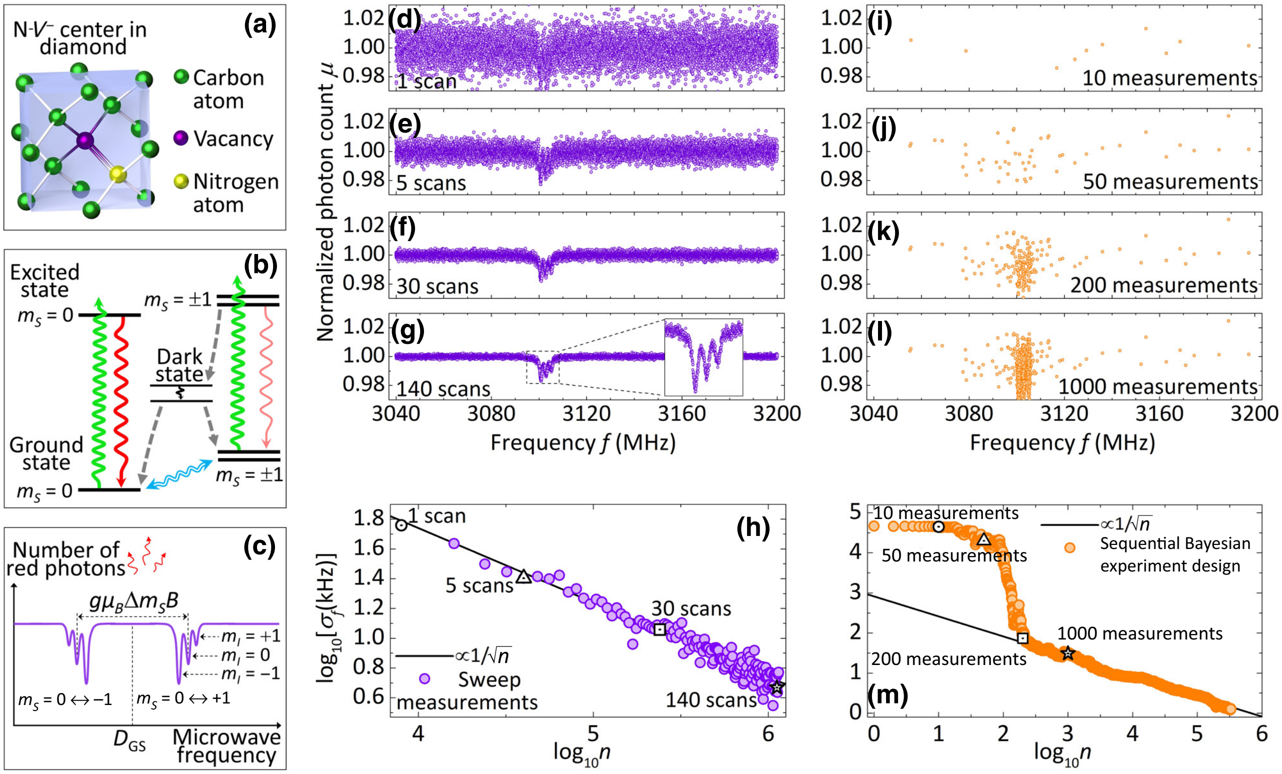

FIG. 1. (a) Crystal structure of the N-V^- center inside a diamond lattice. Green spheres denote carbon atoms, the yellow sphere is a nitrogen atom, and the purple sphere is a vacancy. Each white line corresponds to an sp^3 bond created by a pair of electrons. (b) Schematic structure of the transitions between energy levels of the N-V^- center. The N-V^- center in the ground state can be excited by laser light (green arrows represent transitions due to the absorbed photons); the process preserves spin projection m_S. From the excited state, the N-V^- center can relax back to the ground state by emitting a red photon ($m_S = \pm 1$ or $m_S = 0$ excited states; red arrows represent transitions due to the emitted photons) or can relax nonradiatively through the dark state (only $m_S = \pm 1$ excited states; dashed gray arrow). Transition between the states with $m_S = \pm 1$ and $m_S = 0$ can be induced by microwaves (blue arrow). (c) Schematics of the photoluminescence spectrum of the N-V^- center under application of microwave irradiation and the external magnetic field B. The six dips are present due to the Zeeman splitting and hyperfine interaction. (d)–(g) Averaged data from (d) one scan, (e) five scans, (f) 30 scans, and (g) 140 scans (the inset shows an enlarged signal area) of conventional N-V^--center magnetometry using photoluminescence detection under sweeping of the microwave frequency. The magnetic field is calculated with use of the position of the signal (central dip) in the photoluminescence spectrum. (h) Dependence of the standard deviation σ_f of the signal frequency f_B on the number of photoluminescence measurements n. Each solid purple circle corresponds to a unique number of averaged frequency-sweep scans; each scan consists of 8000 measured data points. Black symbols correspond to the data from (d)–(g). The solid black line shows inverse-square-root scaling. (i)–(l) The data from (i) 10, (j) 50, (k) 200, and (l) 1000 photoluminescence measurements by N-V^--center magnetometry using sequential Bayesian experiment design. (m) Dependence of the standard deviation σ_f of the signal frequency f_B on the number of photoluminescence measurements n. Each solid orange circle corresponds to a unique number of photoluminescence measurements. Black symbols correspond to the data from (i)–(l). The solid black line shows inverse-square-root scaling.

$A_{\mathrm{GS}}^{\mathrm{HF}}/h$ is the hyperfine splitting, a is an overall contrast factor, Ω is a linewidth, and k_{NP} characterizes the nuclear polarization. The coupling between N-V^--center electrons and the nitrogen-nucleus spin (naturally abundant ^{14}N, $I = 1$) leads to the weak spin transfer of constant polarization of the electron spin to the nucleus. However, the nitrogen nucleus is not fully polarized in the presence of slight misalignment of the external magnetic field with the axis of the N-V^- center [49,50]. This leads to the splitting of the N-V^--center transitions into three photoluminescence dips of different amplitudes corresponding to $m_I = -1$, 0, and $+1$, which are separated in frequency by the hyperfine splitting Δf_{HF} [Fig. 1(c)]. For every measurement with microwave excitation, a reference photon count with microwave irradiation switched off is used as a normalizing factor. Throughout this paper, we treat the excitation frequency f as the lone experimental setting design $d = \{f\}$ and the five parameters $\boldsymbol{\theta} = \{f_B,\ \Delta f_{\mathrm{HF}},\ a,\ \Omega,\ k_{\mathrm{NP}}\}$ as unknowns.

We use the triple-resonance spectrum described by Eq. (2) to compare the effectiveness of measurement protocols. The goal of the experiment is to determine the center resonance frequency f_B. The external magnetic field in N-V^{-}-center magnetometry is given by the equation $|B| = (h/g\mu_B)(f_B - D_{\mathrm{GS}})$, where $g\mu_B/h \approx 28$ MHz/mT is the combination of the physical constants. The search range for the signal frequency is from 3040 to 3200 MHz, which corresponds to a magnetic field in the range from 6 to 12 mT. The generated electromagnetic field is set to $B \approx 8.32$ mT (picked by a random-number generator) for the results shown in this paper, corresponding to the N-V^{-}-center resonance frequency $f_B \approx 3103$ MHz. The field is treated as an unknown in the measurements and data analysis.

In the conventional N-V^{-}-center-magnetometry measurements, the photoluminescence of the sample is monitored while the microwave frequency is scanned from 3040 to 3200 MHz with 20-kHz step. Hence, each frequency scan consists of 8000 normalized photoluminescence measurements.

The sequential-Bayesian-experiment-design measurements are iterated over a three-step cycle comprising a setting choice (design) from the allowed microwave frequencies, measurement, and data analysis via Bayesian inference. Here we provide an overview of the process, and direct the interested reader to Secs. S.2 and S.3 in Supplemental Material [44] and Refs. [27,34,51,52] for more-detailed descriptions.

Bayesian methods treat the unknown parameters $\boldsymbol{\theta}$ as random variables with a probability distribution $p(\boldsymbol{\theta})$. In this application, $\boldsymbol{\theta} = \{f_B,\ \Delta f_{\mathrm{HF}},\ a,\ \Omega,\ k_{\mathrm{NP}}\}$ are the parameters of the model function given in Eq. (2). After n iterations, the parameters are described by a conditional distribution $p(\boldsymbol{\theta}|\boldsymbol{y}_n,\ \boldsymbol{d}_n)$ given accumulated measurement results $\boldsymbol{y}_n = \{y_1,\ y_2,\ \ldots,\ y_n\}$ obtained at frequency settings (designs) $\boldsymbol{d}_n = \{d_1, d_2, \ldots,\ d_n\}$.

In the $(n+1)$th iteration, the experiment-design step uses the parameter distribution $p(\boldsymbol{\theta}|\boldsymbol{y}_n,\ \boldsymbol{d}_n)$ to inform the choice of a setting design d_{n+1} for the next measurement. The algorithm models a distribution of measurement predictions for each possible design and then predicts the average improvement in the parameter distribution that would result from the predicted data. "Improvement" is quantified as a predicted change in the information entropy of the parameter distribution and is expressed as a utility function $U(d)$ [53,54]. The derivation of $U(d)$ produces a qualitatively intuitive result: it does the most good to "pin down" the measurement results where they are sensitive to parameter variations. The new setting d_{n+1} is selected to maximize $U(d)$.

After the setting d_{n+1} is used to obtain the measurement result y_{n+1}, these values are used to refine the parameter distribution. With use of Bayesian inference,

$$p(\boldsymbol{\theta}|\boldsymbol{y}_{n+1},\ \boldsymbol{d}_{n+1}) \propto p(y_{n+1}|\boldsymbol{\theta},\ d_{n+1})\,p(\boldsymbol{\theta}|\boldsymbol{y}_n,\ \boldsymbol{d}_n), \qquad (3)$$

where $p(y_{n+1}|\boldsymbol{\theta}, d_{n+1})$ is the *likelihood*, the probability of observing the measured value y_{n+1} calculated for arbitrary parameter values $\boldsymbol{\theta}$ given the frequency setting d_{n+1}. With increasing iteration number, the parameter distribution typically narrows, reflecting increasingly precise estimates of the parameter values.

In each iteration, the sequential-Bayesian-experiment-design algorithm makes an informed setting decision and incorporates new data to inform the next decision. On a qualitative level, the Bayesian method formalizes an intuitive approach of making rough initial measurements to guide later runs, but the Bayesian method offers additional advantages. Bayesian inference incorporates new data, allowing semicontinuous monitoring of "fitting" statistics, and result-based stopping criteria. The utility function provides a nonheuristic, flexible, data-based method for setting decisions. These advantages are especially important for situations where automation is required, speed is essential, or measurement data are expensive.

Software and documentation for sequential Bayesian experiment design are provided online [55].

III. EXPERIMENTAL DETAILS

We used a commercially available single-crystal diamond grown by chemical vapor deposition. The sample size is $3.0 \times 3.0 \times 0.3$ mm^3, with {100} top-surface orientation and surface roughness below 30 nm. The diamond (type IIa) has a nitrogen concentration below 1 ppm and a boron concentration below 0.05 ppm according to the manufacturer. The sample is mounted on top of a 50-mm-long microstrip line, which is used to supply microwaves to manipulate the spin state of the N-V^{-} center. The microstrip line with the sample is placed in an electromagnet between pincer-shaped poles that are oriented to align with the [111] direction of the diamond lattice ($\arcsin\sqrt{2/3} \approx 54.7^{\circ}$ from the vertical). In this arrangement, the magnetic field is pointing along one of the four possible orientations of the N-V^{-}-center axes (vector connecting nitrogen atom to the vacancy site).

A green laser with 520-nm wavelength is used to optically excite the N-V^{-} center. The 0.7-numerical-aperture objective of a custom-built confocal microscope is located above the sample to focus laser excitation inside the diamond and to collect fluorescence from the N-V^{-} center. A dichroic beam splitter with the edge at 650 nm is used to separate excitation laser light from the collected fluorescence. After further wavelength selection with 647-nm long-pass filters, the collected fluorescence is coupled into a multimode fiber and directed to the photon detector. For each data point, a 50-ms photon count with microwave irradiation switched on is divided by a subsequent 50-ms reference count with microwave irradiation switched off. The excitation using green laser light is on continuously. Only 10 mW of microwave power (at the source) and 225 μW of laser power (before the objective) are sent to the sample. The laser power is set by using a linear polarizer and a half-wave plate. The combination of laser power, microwave power, and counting time produces measurements with a signal-to-noise ratio on the order of 1. Such an experimental setup showcases the ability of sequential Bayesian experiment design to locate and measure a complex multiple-peak signal even in extremely noisy data, and shows its broad dynamical range for sensitivity.

IV. RESULTS AND DISCUSSION

First, we report the results of the conventional N-V^--center-magnetometry measurements. Figure 1(d) shows the photoluminescence data measured in one frequency scan. Dips in the photoluminescence spectrum corresponding to optically detected magnetic resonance are visible with a signal-to-noise ratio on the order of 1. We follow the conventional approach to increase the signal-to-noise, which is to remeasure the same scanning range and average the data in the scans. Figures 1(e)– 1(g) show averaged data for increasing numbers of scans. The signal-to-noise ratio increases as the inverse square root of the number of the averaged scans.

To gauge the evolution of parameter uncertainty as a function of scan number, we "fit" the averaged data using Bayesian inference to determine mean values and standard deviations from the parameter distribution. To allow direct comparison, we use the same algorithm for Bayesian inference as in the sequential-design data below. Like the overall signal-to-noise ratio, the standard deviation of the resonance frequency also follows an inverse-square-root dependence on the total number of scans [Fig. 1(h)].

Photoluminescence data from the N-V^--center-magnetometry measurements using sequential Bayesian experiment design are shown in Figs. 1(i)–1(l). Here the data are plotted without averaging. While initial frequency sampling roams across the whole allowed frequency range [Figs. 1(i) and 1(j)], the later measurements almost exclusively focus on the signal location near the resonance dips where the photoluminescence signal is lower [Figs. 1(k) and 1(l)]. The standard deviation σ_f of the center resonance frequency f_B is plotted as a function of the number of measurements in Fig. 1(m). The standard deviation drops by 3 orders of magnitude within the first 200 measurements.

We plot the evolution of the probability distribution $p(\boldsymbol{\theta})$ of the signal frequency f_B and hyperfine splitting $\Delta f_{\rm HF}$ parameters in Fig. 2. The probability distribution is implemented with use of the sequential Monte Carlo method, where the probability density in parameter space is represented by the density of points and by a weight factor attached to each point. After each measurement, the weights are recalculated with use of Bayesian inference. Figure 2(a) shows the initial, *prior* distribution, which consists of 10 000 points distributed through the parameter space with equal weights of 10^{-4} [Fig. 2(a)]. The sum of all weights is normalized to 1.

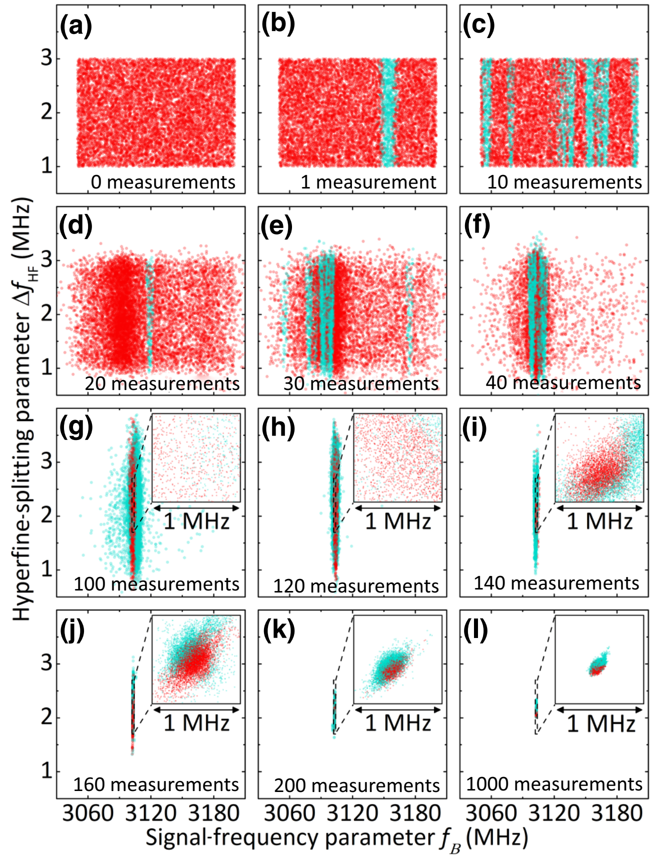

FIG. 2. Dependence of the probability distributions for signal-frequency and hyperfine-splitting parameters on the number of the measurements in N-V^--center magnetometry using sequential Bayesian experiment design. Probability distributions are shown after (a) 0, (b) 1, (c) 10, (d) 20, (e) 30, (f) 40, (g) 100, (h) 120, (i) 140, (j) 160, (k) 200, and (l) 1000 measurements. Each probability distribution consists of 10 000 points in parameter space with weights adding up to 1. Color represents the weight: less than 10^{-4}, cyan; 10^{-4} or greater, red. Insets show enlarged areas of the probability distributions. All insets have the same size (1×1 MHz^2), and span the same parameter space [(3102.5 MHz, 3103.5 MHz); (1.7 MHz, 2.7 MHz)].

Figure 2(b) plots the probability distribution after the first measurement, which yields $\mu_1 = 1.014$ for the normalized photon count at $f_1 = 3154.26$ MHz. Since the resonances are dips in the photon count, values of μ greater than 1 reduce the likelihood that the resonances are located near the measurement frequency f_1. To highlight this effect, distribution points with weights $w < 10^{-4}$ are colored cyan and distribution points with weights $w \geq 10^{-4}$ are colored red. After several cycles of measurements and updating of the weights, a resampling algorithm redistributes points, allowing high-weight points to survive, multiply, and diffuse slightly, while low-weight points face a greater probability of elimination (see Sec. S.4 in Supplemental Material [44]). Resampling allows the computational resources to be focused on high-probability regions of parameter space without completely abandoning low-probability regions. The effects of resampling are visible in Figs. 2(d)–2(l) as a growing density of points near 3100 MHz. After the first 200 measurements, the $p(f_B)$ distribution has effectively contracted from spanning a range of 150 MHz to less than 1 MHz [Figs. 2(k) and 2(l)]. Redistribution of the weights also allows the probability distribution to diffuse beyond the initial boundary conditions. For example, initial weights occupy the Δf_{HF} parameter space from 1 to 3 MHz [Figs. 2(a)–2(c)], but after 100 measurements the resampling steps have allowed the probability distribution to span the Δf_{HF} parameter space from 0.5 to 4 MHz. This diffusion allows slow convergence to values outside the prior distribution (i.e., in the areas where the experimenter does not expect to find final parameters' values), which is helpful in cases when the experimenter does not have an accurate initial estimate for the parameter.

The evolution of the N-V^{-}-center-magnetometry measurements using sequential Bayesian experiment design is in sharp contrast with the evolution of the conventional N-V^{-}-center-magnetometry measurements. The standard deviation of the signal frequency using sequential Bayesian experiment design follows the typical pattern displayed in Fig. 1(m). After an initial period of broad sampling of the parameter space, the algorithm focuses measurements near the resonance frequencies [Fig. 3(a)] and the probability distribution $p(f_B)$ contracts rapidly. After this contraction, the standard deviation of f_B decreases as the inverse square root of the total number of measurements n [Fig. 1(m)]. In contrast, the standard deviation of the signal frequency in the swept-frequency measurements does not go through such rapid contraction phase and follows an inverse-square-root-of-n scaling from the beginning [Fig. 1(h)].

The difference in the measurement strategies can be clearly seen in the photoluminescence data for the first 1000 measurements. Sequential Bayesian experiment design has already narrowed down the probability distribution $p(f_B)$ for the signal frequency, and most of

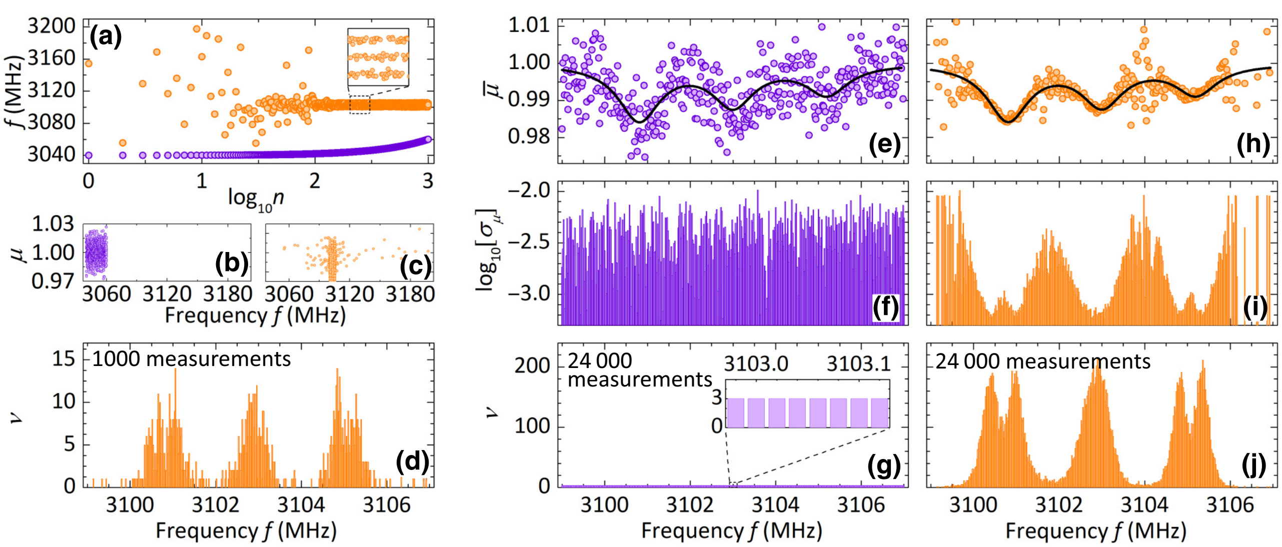

FIG. 3. (a) Dependence of the measurement frequency on the number of measurements for the conventional-N-V^{-}-center-magnetometry microwave-frequency-sweep scan (solid purple circles) and N-V^{-}-center magnetometry using sequential Bayesian experiment design (solid orange circles). The inset shows an enlarged view of the area enclosed by the dashed rectangle. Photoluminescence data for the first 1000 measurements from (b) the conventional N-V^{-}-center-magnetometry microwave-frequency-sweep scan and (c) N-V^{-}-center magnetometry using sequential Bayesian experiment design. (d) Distribution of the measurement frequency for the first 1000 measurements by N-V^{-}-center magnetometry using sequential Bayesian experiment design. Dependence of (e),(h) the average normalized photon count $\bar{\mu}$, (f),(i) logarithm of the standard deviation of the normalized photon count σ_μ, and (g),(j) number of measurements $\nu(f)$ on the measurement frequency for the first 24 000 measurements. (e)–(g) Data from the conventional-N-V^{-}-center magnetometry scan (purple); (h)–(j) data from N-V^{-}-center magnetometry using sequential Bayesian experiment design (orange). The solid black line in (e),(h) shows fitting using the function μ of *all* the measured data: 140 scans (1 120 000 measurements) of conventional N-V^{-}-center magnetometry and 330 000 measurements by N-V^{-}-center magnetometry using sequential Bayesian experiment design. The inset in (g) provides an enlarged view of the data.

the measurements are taken at the signal position—the location of the three hyperfine-split dips [Fig. 3(a) orange solid circles and Figs. 3(c) and 3(d)]. In contrast, the frequency sweep in the conventional measurements has not even reached the frequency where the signal is located, and all 1000 data points are spent on measuring the background [Fig. 3(a) purple solid circles and Fig. 3(b)]. After 24 000 measurements (three full-range conventional sweep scans), only three measurements are performed at each frequency at the signal location by the conventional N-V^{-}-center magnetometry [Fig. 3(g)], compared with a peak of 214 measurements per frequency for sequential-Bayesian-experiment-design measurements [Fig. 3(j)]. This concentration of measurements results in a standard deviation of the averaged Bayesian measurement [Fig 3(i)] that is an order of magnitude smaller than in the conventional measurement [Fig. 3(f)].

An interesting behavior of the utility function $U(\boldsymbol{d}=\{f\})$ can be seen in Fig. 3(j). In the central, $m_I = 0$ photoluminescence-dip area, most of the measurements are concentrated near its center (frequency f_B), while at the outer dips located at $f_B - \Delta f_{\rm HF}$ and $f_B + \Delta f_{\rm HF}$, measurements are concentrated on the sides of the dips, producing double-peak structures in the distribution of the measurements [Fig. 3(j)]. In simulations and measurements on single-dip resonances, similar focus on the sides of dips is typical behavior, and it is consistent with the high sensitivity of the sides of the dip model to the resonance-frequency parameter. On the other hand, the central concentration of measurements that we observe at the central dip in Fig. 3(j) would be atypical behavior for single resonances. We speculate that this behavior stems from the triple-resonance model's [Eq. (2)] implicit assumption that the center resonance lies at the midpoint between the outer resonances.

The "smart" measurement strategy of taking data into account on the fly—instead of waiting until the end of the experiment—allows N-V^{-}-center magnetometry based on sequential Bayesian experiment design to dramatically outperform conventional N-V^{-}-center magnetometry. For example, to achieve the precision of $\sigma_f = 5.5 \times 10^{-3}$ MHz standard deviation of the signal frequency, the conventional sweep-based N-V^{-}-center magnetometry requires 10^6 measurements, while N-V^{-}-center magnetometry based on sequential Bayesian experiment design requires only 24 350 measurements to achieve the same precision. Using the ratio between $1/\sqrt{n}$ scaling of the standard deviations of the signal frequencies for two methods (Fig. 4), we determine sequential-Bayesian-experiment-design magnetometry to be 45 times faster than the conventional measurement approach.

Up to this point, we have compared measurement protocols on the basis of the number of measurements, but "wall-clock" time may be a more relevant basis for comparison, since sequential Bayesian experiment design

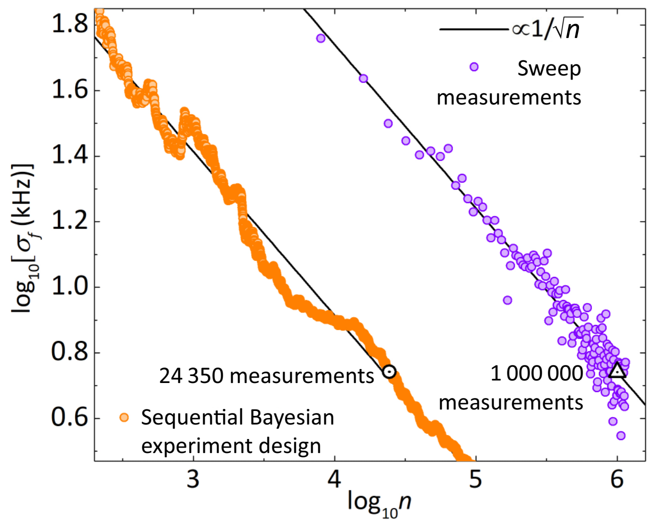

FIG. 4. Dependence of the standard deviation of the signal frequency σ_f on the number of photoluminescence measurements. Each filled orange circle corresponds to a unique number of photoluminescence measurements using sequential Bayesian experiment design. Each filled purple circle corresponds to a unique number of averaged frequency-sweep scans; each scan consists of 8000 measured photoluminescence data points. Black symbols correspond to equal standard deviation of the signal frequency for sequential Bayesian experiment design (black circle) and conventional sweep measurement (black triangle). Solid black lines show inverse-square-root scaling.

comes with an added cost of computational time. Photons from N-V^{-} centers are counted for 100 ms at each data point (50 ms with microwave irradiation switched on, followed by 50 ms with microwave irradiation switched off). In the conventional protocol, the average time spent measuring one data point is 150 ms. The additional time of 50 ms is spent on communication between the devices, saving data, etc. With sequential Bayesian experiment design, the average time spent measuring one data point is 204 ms, a 36% (54-ms) increase in measurement time compared with the conventional setup. The additional time represents the added computational cost of Bayesian inference and utility calculations for each measurement. The computation time depends on the computer hardware and programming methods. Here we report results using a single core of a processor of an ordinary PC programmed in PYTHON using the NUMPY package (see Sec. S.4 in Supplemental Material [44]). Compiled code and parallel computation offer avenues for significant reductions in computation time [56,57]. The cost of an additional processor (several hundred dollars) is also negligible compared with the cost of the other hardware typically used in N-V^{-}-center-magnetometry experiments. However, in the light of the 4400% speedup, the associated additional Bayesian computation time (36% longer measurement time) is negligible, even when the ordinary processor is used and without use of parallel threads.

In the N-V^--center measurements that we perform using sequential Bayesian experiment design, we always observe speedup of more than 1 order of magnitude. The amount of speedup depends on the experimental setup, signal, set of parameters and settings, and reaches close to 2 orders of magnitude for some of the experiments that we perform. A big factor that influences the speedup is the fraction of the settings space occupied by the signal, compared with the whole space spanned by the settings d (scanning or sensing range). In the experiment described in this paper, the signal occupies roughly 10% of the whole scanning range (16 MHz out of the 160-MHz frequency range: 8 MHz is occupied by the dips and 4 MHz is occupied on each side by their shoulders). This value can be much smaller in magnetometers or sensors with a broad sensing range, which will lead to even larger speedups. However, a focus on the measurements with maximum utility function allows sequential Bayesian experiment design to be beneficial even for measurements where a signal is present throughout the whole settings space d (see Sec. S.5 in Supplemental Material for more details [44]). As a rule of thumb, the more time an experimental procedure spends on measuring data with low utility-function values (e.g., areas away from the signal or areas with a small signal-to-noise ratio), the more beneficial will be implementation of measurements using sequential Bayesian experiment design. Sequential Bayesian experiment design can be particularly useful for maturing N-V^--center-magnetometry technology and moving it into the market. Scanning magnetometers or compact in-the-field sensors need to obtain data as fast as possible. Sequential Bayesian experiment design can be used as a much-faster alternative to the numerous averaging scans. It can also be combined with other approaches that increase sensitivity, such as magnetometry using complicated pulse sequences. While the current study focuses on N-V^--center magnetometry using sequential Bayesian experiment design, the reported methods—and corresponding speedups—are directly transferable to other areas beyond N-V^--center magnetometry.

V. CONCLUSION

We report a speedup of more than 1 order of magnitude of N-V-center magnetometry using sequential Bayesian experimental design compared with conventional N-V^--center magnetometry. The large gain in the speed and precision of N-V^--center magnetometry using sequential Bayesian experiment design demonstrated in this study is readily translatable to other applications beyond magnetometry and experiments with N-V^- centers. The software (optbayesexpt) developed to perform sequential-Bayesian-experiment-design measurements is available online for public use free of charge.

ACKNOWLEDGMENTS

S.D. and K.A. acknowledge support under the cooperative research agreement between the University of Maryland and the National Institute of Standards and Technology Physical Measurement Laboratory (Grant No. 70NANB14H209) through the University of Maryland. The authors thank Adam Pintar for many helpful discussions.

[1] Y. Wu, Y. Wang, X. Qin, X. Rong, and J. Du, A programmable two-qubit solid-state quantum processor under ambient conditions, NPJ Quantum Inf. **5**, 9 (2019).

[2] A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J.-P. Poizat, and P. Grangier, Single Photon Quantum Cryptography, Phys. Rev. Lett. **89**, 187901 (2002).

[3] G. D. Fuchs, G. Burkard, P. V. Klimov, and D. D. Awschalom, A quantum memory intrinsic to single nitrogen–vacancy centres in diamond, Nat. Phys. **7**, 789 (2011).

[4] W. L. Yang, Z. Q. Yin, Y. Hu, M. Feng, and J. F. Du, High-fidelity quantum memory using nitrogen-vacancy center ensemble for hybrid quantum computation, Phys. Rev. A **84**, 010301 (2011).

[5] V. Vaijayanthimala and H.-C. Chang, Functionalized fluorescent nanodiamonds for biomedical applications, Nanomedicine **4**, 47 (2009).

[6] I. Badea and R. Kaur, Nanodiamonds as novel nanomaterials for biomedical applications: Drug delivery and imaging systems, Int. J. Nanomedicine **8**, 203 (2013).

[7] S. Kolkowitz, A. C. Bleszynski Jayich, Q. P. Unterreithmeier, S. D. Bennett, P. Rabl, J. G. E. Harris, and M. D. Lukin, Coherent sensing of a mechanical resonator with a single-spin qubit, Science **335**, 1603 (2012).

[8] G. Kucsko, P. C. Maurer, N. Y. Yao, M. Kubo, H. J. Noh, P. K. Lo, H. Park, and M. D. Lukin, Nanometre-scale thermometry in a living cell, Nature **500**, 54 (2013).

[9] P. Neumann, I. Jakobi, F. Dolde, C. Burk, R. Reuter, G. Waldherr, J. Honert, T. Wolf, A. Brunner, J. H. Shim, D. Suter, H. Sumiya, J. Isoya, and J. Wrachtrup, High-precision nanoscale temperature sensing using single defects in diamond, Nano Lett. **13**, 2738 (2013).

[10] F. Dolde, H. Fedder, M. W. Doherty, T. Nöbauer, F. Rempp, G. Balasubramanian, T. Wolf, F. Reinhard, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup, Electric-field sensing using single diamond spins, Nat. Phys. **7**, 459 (2011).

[11] G. Balasubramanian, I. Y. Chan, R. Kolesov, M. Al-Hmoud, J. Tisler, C. Shin, C. Kim, A. Wojcik, P. R. Hemmer, A. Krueger, T. Hanke, A. Leitenstorfer, R. Bratschitsch, F. Jelezko, and J. Wrachtrup, Nanoscale imaging magnetometry with diamond spins under ambient conditions, Nature **455**, 648 (2008).

[12] 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 **455**, 644 (2008).

[13] C. Du, T. van der Sar, T. X. Zhou, P. Upadhyaya, F. Casola, H. Zhang, M. C. Onbasli, C. A. Ross, R. L. Walsworth, Y. Tserkovnyak, and A. Yacoby, Control and local measurement of the spin chemical potential in a magnetic insulator, Science **357**, 195 (2017).

[14] 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).

[15] L. Rondin, J.-P. Tetienne, T. Hingant, J.-F. Roch, P. Maletinsky, and V. Jacques, Magnetometry with nitrogen-vacancy defects in diamond, reports prog, Phys. **77**, 056503 (2014).

[16] P. Maletinsky, S. Hong, M. S. Grinolds, B. Hausmann, M. D. Lukin, R. L. Walsworth, M. Loncar, and A. Yacoby, A robust scanning diamond sensor for nanoscale imaging with single nitrogen-vacancy centres, Nat. Nanotechnol. **7**, 320 (2012).

[17] P. Reith and H. Hilgenkamp, Analysing magnetism using scanning squid microscopy, Rev. Sci. Instrum. **88**, 123706 (2017).

[18] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[19] M. W. Mitchell and S. Palacios Alvarez, ColloquiumâĂŕ: Quantum limits to the energy resolution of magnetic field sensors, Rev. Mod. Phys. **92**, 021001 (2020).

[20] J. M. Schloss, J. F. Barry, M. J. Turner, and R. L. Walsworth, Simultaneous Broadband Vector Magnetometry Using Solid-State Spins, Phys. Rev. Appl. **10**, 034044 (2018).

[21] P. Balasubramanian, C. Osterkamp, Y. Chen, X. Chen, T. Teraji, E. Wu, B. Naydenov, and F. Jelezko, Dc magnetometry with engineered nitrogen-vacancy spin ensembles in diamond, Nano Lett. **19**, 6681 (2019).

[22] C. S. Shin, C. E. Avalos, M. C. Butler, D. R. Trease, S. J. Seltzer, J. Peter Mustonen, D. J. Kennedy, V. M. Acosta, D. Budker, A. Pines, and V. S. Bajaj, Room-temperature operation of a radiofrequency diamond magnetometer near the shot-noise limit, J. Appl. Phys. **112**, 124519 (2012).

[23] H. A. R. El-Ella, S. Ahmadi, A. M. Wojciechowski, A. Huck, and U. L. Andersen, Optimised frequency modulation for continuous-wave optical magnetic resonance sensing using nitrogen-vacancy ensembles, Opt. Express **25**, 14809 (2017).

[24] H. Clevenson, L. M. Pham, C. Teale, K. Johnson, D. Englund, and D. Braje, Robust high-dynamic-range vector magnetometry with nitrogen-vacancy centers in diamond, Appl. Phys. Lett. **112**, 252406 (2018).

[25] K. Ambal and R. D. McMichael, A differential rate meter for real-time peak tracking in optically detected magnetic resonance at low photon count rates, Rev. Sci. Instrum. **90**, 023907 (2019).

[26] G. Liu, M. Chen, Y.-X. Liu, D. Layden, and P. Cappellaro, Repetitive readout enhanced by machine learning, Mach. Learn. Sci. Technol. **1**, 015003 (2020).

[27] K. Chaloner and I. Verdinelli, Bayesian experimental design: A review, Stat. Sci. **10**, 273 (1995).

[28] T. Bayes, An essay towards solving a problem in the doctrine of chances, Philos. Trans. R. Soc. London **53**, 370 (1763).

[29] P. S. Laplace, Memoire sur la probabilite des causes par les evenemens, Mem. Math. Phys. **6**, 621 (1774).

[30] P. S. Laplace, Memoir on the probability of the causes of events, Stat. Sci. **1**, 364 (1986).

[31] E. T. Jaynes, *Probability Theory* (Cambridge University Press, Cambridge, 2003).

[32] H. Mabuchi, Dynamical identification of open quantum systems, Quantum Semiclass. Opt. J. Eur. Opt. Soc. Part B **8**, 1103 (1996).

[33] J. Gambetta and H. M. Wiseman, State and dynamical parameter estimation for open quantum systems, Phys. Rev. A **64**, 042105 (2001).

[34] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, Robust online Hamiltonian learning, New J. Phys. **14**, 103013 (2012).

[35] S. Gammelmark and K. Mølmer, Bayesian parameter inference from continuously monitored quantum systems, Phys. Rev. A **87**, 032115 (2013).

[36] E. Scerri, E. M. Gauger, and C. Bonato, Extending qubit coherence by adaptive quantum environment learning, New J. Phys. **22**, 035002 (2020).

[37] A. Negretti and K. Mølmer, Estimation of classical parameters via continuous probing of complementary quantum observables, New J. Phys. **15**, 125002 (2013).

[38] H. T. Dinani, D. W. Berry, R. Gonzalez, J. R. Maze, and C. Bonato, Bayesian estimation for quantum sensing in the absence of single-shot detection, Phys. Rev. B **99**, 125413 (2019).

[39] I. Schwartz, J. Rosskopf, S. Schmitt, B. Tratzmiller, Q. Chen, L. P. McGuinness, F. Jelezko, and M. B. Plenio, Blueprint for nanoscale nmr, Sci. Rep. **9**, 6938 (2019).

[40] C. Bonato and D. W. Berry, Adaptive tracking of a time-varying field with a quantum sensor, Phys. Rev. A **95**, 1 (2017).

[41] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'brien, J. G. Rarity, A. Laing, and M. G. Thompson, Experimental quantum Hamiltonian learning, Nat. Phys. **13**, 551 (2017).

[42] C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, Optimized quantum sensing with a single electron spin using real-time adaptive measurements, Nat. Nanotechnol. **11**, 247 (2016).

[43] R. Santagati, A. A. Gentile, S. Knauer, S. Schmitt, S. Paesani, C. Granade, N. Wiebe, C. Osterkamp, L. P. McGuinness, J. Wang, M. G. Thompson, J. G. Rarity, F. Jelezko, and A. Laing, Magnetic-field learning using a single electronic spin in diamond with one-photon readout at room temperature, Phys. Rev. X **9**, 021019 (2019).

[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.14.054036 for additional details and discussion on the structure and physics of the N-V center, sequential Bayesian experiment design, implementation of probability distributions, specifications of the computational hardware used for sequential Bayesian experiment design, and speedup of the sequential Bayesian experiment design.

[45] A. Gruber, A. Drabenstedt, C. Tietz, L. Fleury, J. Wrachtrup, and C. von Borczyskowski, Scanning confocal

optical microscopy and magnetic resonance on single defect centers, Science **276**, 2012 (1997).
[46] G. Davies and M. F. Hamer, Optical studies of the 1.945 eV vibronic band in diamond, Proc. R. Soc. London. A. Math. Phys. Sci. **348**, 285 (1976).
[47] J. Kohler, J. A. J. M. Disselhorst, M. C. J. M. Donckers, E. J. J. Groenen, J. Schmidt, and W. E. Moerner, Magnetic resonance of a single molecular spin, Nature **363**, 242 (1993).
[48] J. Wrachtrup, C. von Borczyskowski, J. Bernard, M. Orrit, and R. Brown, Optical detection of magnetic resonance in a single molecule, Nature **363**, 244 (1993).
[49] V. Jacques, P. Neumann, J. Beck, M. Markham, D. Twitchen, J. Meijer, F. Kaiser, G. Balasubramanian, F. Jelezko, and J. Wrachtrup, Dynamic Polarization of Single Nuclear Spins by Optical Pumping of Nitrogen-Vacancy Color Centers in Diamond at Room Temperature, Phys. Rev. Lett. **102**, 057403 (2009).
[50] R. Fischer, A. Jarmola, P. Kehayias, and D. Budker, Optical polarization of nuclear ensembles in diamond, Phys. Rev. B **87**, 125207 (2013).
[51] D. V. Lindley, On a measure of the information provided by an experiment, Ann. Math. Stat. **27**, 986 (1956).
[52] X. Huan and Y. M. Marzouk, Simulation-based optimal Bayesian experimental design for nonlinear systems, J. Comput. Phys. **232**, 288 (2013).
[53] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. **22**, 79 (1951).
[54] S. Kullback, *Information Theory and Statistics* (Dover Publications, Mineola, 1968).
[55] R.D. McMichael, Optimal Bayesian Experiment Design Software [Online] (2020). https://github.com/usnistgov/optbayesexpt/.R.D. McMichael, Optimal Bayesian Experiment Design Documentation [Online] (2020). https://pages.nist.gov/optbayesexpt/.
[56] E. G. Ryan, C. C. Drovandi, J. M. McGree, and A. N. Pettitt, A review of modern computational algorithms for Bayesian optimal design, Int. Stat. Rev. **84**, 128 (2016).
[57] M. A. Nicely and B. E. Wells, Improved parallel resampling methods for particle filtering, IEEE Access **7**, 47593 (2019).