

## Real-time adaptive estimation of decoherence timescales for a single qubit

Muhammad Junaid Arshad<sup>1</sup>, Christiaan Bekker<sup>1</sup>, Ben Haylock<sup>1</sup>, Krzysztof Skrzypczak<sup>1</sup>, Daniel White<sup>1</sup>, Benjamin Griffiths<sup>2</sup>, Joe Gore<sup>3</sup>, Gavin W. Morley<sup>3</sup>, Patrick Salter<sup>4</sup>, Jason Smith<sup>2</sup>, Inbar Zohar<sup>5</sup>, Amit Finkler<sup>5</sup>, Yoann Altmann<sup>6</sup>, Erik M. Gauger<sup>1</sup> and Cristian Bonato<sup>1,\*</sup>

<sup>1</sup>*SUPA, Institute of Photonics and Quantum Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom*


<sup>2</sup>*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*

<sup>3</sup>*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*

<sup>4</sup>*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, United Kingdom*

<sup>5</sup>*Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 7610001, Israel*

<sup>6</sup>*Institute of Signals, Sensors and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom*

 (Received 25 October 2023; revised 16 January 2024; accepted 18 January 2024; published 13 February 2024)

Characterizing the time over which quantum coherence survives is critical for any implementation of quantum bits, memories, and sensors. The usual method for determining a quantum system's decoherence rate involves a suite of experiments probing the entire expected range of this parameter, and extracting the resulting estimation in postprocessing. Here we present an adaptive multiparameter Bayesian approach, based on a simple analytical update rule, to estimate the key decoherence timescales ( $T_1$ ,  $T_2^*$ , and  $T_2$ ) and the corresponding decay exponent of a quantum system in real time, using information gained in preceding experiments. This approach reduces the time required to reach a given uncertainty by a factor up to an order of magnitude, depending on the specific experiment, compared to the standard protocol of curve fitting. A further speedup of a factor approximately 2 can be realized by performing our optimization with respect to sensitivity as opposed to variance.

DOI: [10.1103/PhysRevApplied.21.024026](https://doi.org/10.1103/PhysRevApplied.21.024026)

### I. INTRODUCTION

Decoherence, resulting from the interaction of a quantum system with its environment, is a key performance indicator for qubits in quantum technologies [1], including quantum communication, computation, and sensing. Decoherence timescales determine the storage time for quantum memories and quantum repeaters, a crucial metric for quantum communication networks [2–5]. Rapid benchmarking of decoherence timescales in platforms such as superconducting qubits [6,7] or silicon spin qubits [8,9], is a critical validation and quality assurance step for the development of large-scale quantum computing architectures, and has the potential to improve error-correction protocols efficiently close to fault-tolerance thresholds. In quantum sensing, the role of decoherence is twofold. On

one side, decoherence sets the ultimate performance limit of the sensors [10]. On the other hand, decoherence itself can be the quantity measured by a quantum sensor, as it provides information about the environment. An example of this is relaxometry, where the rate at which a polarized quantum sensor reaches the thermal equilibrium configuration gives information about different physical processes in the environment [11–14].

Decoherence rates can be measured by preparing the system into a known quantum state and probing it at varying time delays to determine the probability of decay from its initial state. The standard protocol for decoherence estimation involves a series of measurements with time delays set over a predetermined range, reflecting the expected value of the decoherence rate, and fitting of the result to a decay function. As the range of time delays is determined in advance, some of the measurements will provide little information on the decoherence of the system, since the time delay is either much shorter than the true decoherence rate, resulting in no decay, or much longer, resulting in complete decay.

Here we introduce a real-time adaptive protocol to measure decoherence timescales  $T_1$ ,  $T_2^*$  and  $T_2$  for a

\*c.bonato@hw.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

single qubit [15], respectively corresponding to relaxation, dephasing, and echo decay time [1], together with the coherence decay exponent  $\beta$ . While the proposed algorithms are very general and can be applied to any quantum architecture, our experiments are implemented on a single spin qubit associated with a nitrogen-vacancy (N-V) center in diamond.

Adaptive techniques have been shown to be central to progress across a broad range of quantum technologies [16]. Early work in this field involved the implementation of adaptive quantum phase-estimation algorithms on photonic systems [17], later extended to frequency estimation with applications to (static) dc magnetometry with single electron spins [18–20]. Alternative adaptive protocols for the estimation of static magnetic fields are based on sequential Bayesian experiment design [21,22] and ad hoc heuristics [23], later applied to the characterization of a single nuclear spin [24]. Real-time adaptation of experimental settings has also been shown to be advantageous when measuring spin relaxation [25] or tracking the magnetic resonance of a single electron spin in real time [26–29]. Furthermore, adaptive techniques have been investigated to enhance photonic quantum sensors [30–32], as a control tool for quantum state preparation [33] and to extend quantum coherence of a qubit by manipulating the environment [34–36].

Despite these pioneering experiments, several methodological questions still remain open. A priority concern is that adaptive protocols introduce an overhead, given by the time required to compute settings on the fly for the next iteration. It is crucial to minimize this computation time, since it can slow the protocol down to the point that the overhead can reverse the gain in measurement speed compared to a simple parameter sweep. This has not been considered in many cases, in particular where algorithms were investigated through computer simulations [21,22,27,37] or as offline processing of pre-existing experimental data [23]. While the optimization of complex utility functions can possibly deliver the best theoretical results, this could be practically less advantageous than near-optimal approaches with very fast update rules in minimizing total measurement duration. A second issue is related to the fact that, for multiparameter Hamiltonian estimation, standard approaches such as the maximization of Fisher information can fail, as the Fisher-information matrix becomes singular when controlling the evolution time [38]. This has stimulated researchers to find ad hoc heuristics, for example, the particle guess heuristic [23,24,38] for the estimation of Hamiltonian terms; these heuristics, however, do not necessarily work beyond Hamiltonian estimation. A third question is related to what quantity should be optimized. Previous work has targeted the minimization of the variance of the probability distribution for the quantity of interest [24,39]. While this is clear when all measurements feature the same duration,

the answer is less straightforward when adapting the probing time. If two measurements with different probing times result in a similar variance, the protocol should prefer the shorter one, minimizing the overall sensing time.

Here we address these open questions, presenting theoretical and experimental data about the adaptive estimation of decoherence for a single qubit, using N-V centers as a case study. Compared to other recent investigations of adaptive protocols [23–25], our experiments utilize a very simple analytical update rule based on the concept of Fisher information and the Cramér-Rao bound. By exploiting state-of-the-art fast electronics, we experimentally perform the real-time processing in less than 50  $\mu$ s, an order of magnitude shorter than previous real-time experiments [24], negligible compared to the duration of each measurement. Such a short timescale makes our approach useful for qubits where fast single-shot readout is available such as trapped ions [40], superconducting qubits [41], and several types of spin qubits [42–45], and could be further shortened in future work by implementing the protocols on field-programmable gate array (FPGA) hardware.

In the case of multiparameter estimation, previous work on Hamiltonian estimation had pointed out that the Cramér-Rao bound cannot be used in the optimization as the Fisher-information matrix is singular and cannot be inverted [38]. Here we address this issue by utilizing multiple probing times, showing that the Fisher-information matrix can be inverted and that the corresponding adaptive scheme provides better performance than nonadaptive approaches. Finally, we discuss what quantity needs to be targeted to achieve the best sensor performance, experimentally demonstrating the superiority of optimizing sensitivity, defined as variance multiplied by time, over optimizing variance. As a figure of merit, sensitivity encourages faster measurements.

Our work tackles these general questions using the characterization of decoherence as a test case. While adaptive approaches have been investigated in the case of phase and frequency estimation [17–24], also in relation to Hamiltonian learning [23], the case of decoherence is much less explored, with only one work targeting the estimation of the relaxation timescale  $T_1$  [25]. Here we provide the first complete characterization of the three decoherence timescales typically used in experiments ( $T_1$ ,  $T_2^*$ , and  $T_2$ ), together with the decoherence decay exponent  $\beta$ .

## II. THEORY

Decoherence and relaxation are processes induced by the interaction of a qubit with its environment, leading to random transitions between states or random-phase accumulation during the evolution of the qubit. These processes are typically estimated by preparing a quantum state and tracking the probability of still measuring the initial state

over time, which can be captured by a functional form [10]

$$p(t) \propto \frac{1}{2} (1 - e^{-\chi(t)}) \quad . \quad (1)$$

Although the noise processes induced by interaction with the environment can be complex,  $\chi(t)$  can often be approximated by a simple power law:

$$\chi(t) \propto \left(\frac{t}{T_\chi}\right)^\beta, \quad (2)$$

where  $T_\chi$  and  $\beta$  depend on the specific noise process [1]. For white noise, the decay is exponential with  $\beta = 1$ . For a generic  $1/f^q$  decay, relevant, for example, for superconducting qubits, with a noise spectral density as  $\propto \omega^q$ ,  $\chi(t)$  scales as  $\chi(t) \propto (t/T_\chi)^{1+q}$  [46].

In the case of a single electronic spin dipolarly coupled to a diluted bath of nuclear spins, the decay exponents have been thoroughly investigated, with analytical solutions available for different parameter regimes [47]. If the intrabath coupling can be neglected, the free induction decay of a single spin is approximately Gaussian ( $\beta = 2$ ) [48,49]. The Hahn-echo decay exponent  $T_2$  can vary, typically between  $\beta \sim 1.5 - 4$  depending on the specific bath parameters and applied static magnetic field [47,50].

In Sec. II A, we assume the decay exponent  $\beta$  to be known, and we focus only on estimating the decay time  $T_\chi$ . This is a practically relevant situation in cases where the nature of the bath is well understood and the decay exponent  $\beta$  is known, at least approximately, *a priori*. We then extend our analysis to the simultaneous estimation of  $T_\chi$  and  $\beta$  in Sec. II B.

Figure 1(a) sketches the operation of the real-time adaptive sensing system developed in our study. We have utilized the electronic spin associated with a N-V center in diamond as the qubit, which is initialized and readout by optical pulses. The qubit state is manipulated by microwave (MW) pulses, created in real-time by an arbitrary waveform generator (AWG) based on an external digital input. After the application of a pulse sequence, the qubit is optically readout, with the spin-state information encoded in the number of photoluminescence photon counts during optical illumination. The core of our adaptive system is a real-time microcontroller, which uses the detected photon count rate to estimate the values of the decoherence timescale ( $T_\chi$ ) and the decay exponent ( $\beta$ ) via Bayesian inference. As shown in the inset, the probability distribution starts out as uniformly flat, but begins to converge around the true value after a few iterations. Based on the estimated value in the current iteration, the microcontroller computes the optimal probing time ( $\tau$ ) for the subsequent measurement and communicates this value to the AWG, which then constructs the next estimation

sequence accordingly. This cycle repeats for several iterations until a desired level of error in the estimation of the target quantity is reached. Figure 1(b) shows the flow of the experimental estimation sequence. For our experiments, a single measurement of the qubit lacks the information required for discriminating its state effectively. Therefore, we conduct  $R$  measurements, to obtain  $r$  detected photons, enough to discriminate the spin state. Such counts are then utilized to update the probability distribution  $p(T_\chi)$  using Bayes' rule. After the Bayesian update, an updated probability distribution is used to compute the optimal settings and provide feedback for the subsequent measurements. Figure 1(c) shows an example of experimental estimation of  $T_2^*$ , performed by an adaptive Ramsey experiment, plotted as the evolution of  $p(T_2^*)$  for increasing estimation epochs. In the beginning,  $p(T_2^*)$  is a uniform distribution in the range 0–8  $\mu$ s, which then converges to a singly peaked distribution after more and more measurement outcomes are processed. In the case of an N-V center in high-purity diamond, the decay is expected to be Gaussian ( $\beta = 2$ ) [49]. As described later in Sec. II A, the optimal adaptive rule for this case is to choose the probing time as  $\tau_{\text{opt}} \sim 0.89 \times \hat{T}_2^*$  [see Eq. (A5)], where  $\hat{T}_2^*$  is the current estimate of  $T_2^*$  computed from the probability distribution  $p(T_2^*)$ . The chosen values for  $\tau$  are shown on the bottom plot, illustrating how they converge very fast to the optimal value  $\tau_{\text{opt}} \sim 0.89 \times (T_2^*)_{\text{true}} \sim 2.23 \mu$ s.

### A. Adaptive Bayesian estimation

We utilize Bayesian inference, exploiting Bayes' theorem to update knowledge about the decoherence time  $T_\chi$  and decay exponent  $\beta$  in the light of a set of alternative measurement outcomes denoted by  $\vec{m} = \{m_1, m_2, \dots\}$ . Thanks to its flexibility in accounting for experimental imperfections and for integrating real-time adaptation of the experimental setting while remaining easy to interpret mathematically, the Bayesian framework [16,51] has been widely applied in quantum technology, from sensing [19,23–25], to the tuning of quantum circuits [52,53], and model learning [54]. In this section, we will restrict the discussion to the characterization of the decoherence time  $T_\chi$ ; the extension to a multiparameter case, with the simultaneous estimation of  $T_\chi$  and  $\beta$ , will be presented in Sec. II B.

For each binary measurement outcome  $m_n$  ( $m_n = 0, 1$ ), the probability distribution of  $T_\chi$ , which represents our knowledge about  $T_\chi$ , is updated as

$$P(T_\chi | m_{1:n}) \propto P(m_n | T_\chi) P(T_\chi | m_{1:(n-1)}), \quad (3)$$

where  $m_{1:n} = \{m_1, \dots, m_n\}$ . Here,  $P(T_\chi | m_{1:(n-1)})$  is the posterior probability after the  $(n-1)$ th update and proceeds to serve as prior distribution at the  $n$ th iteration, and

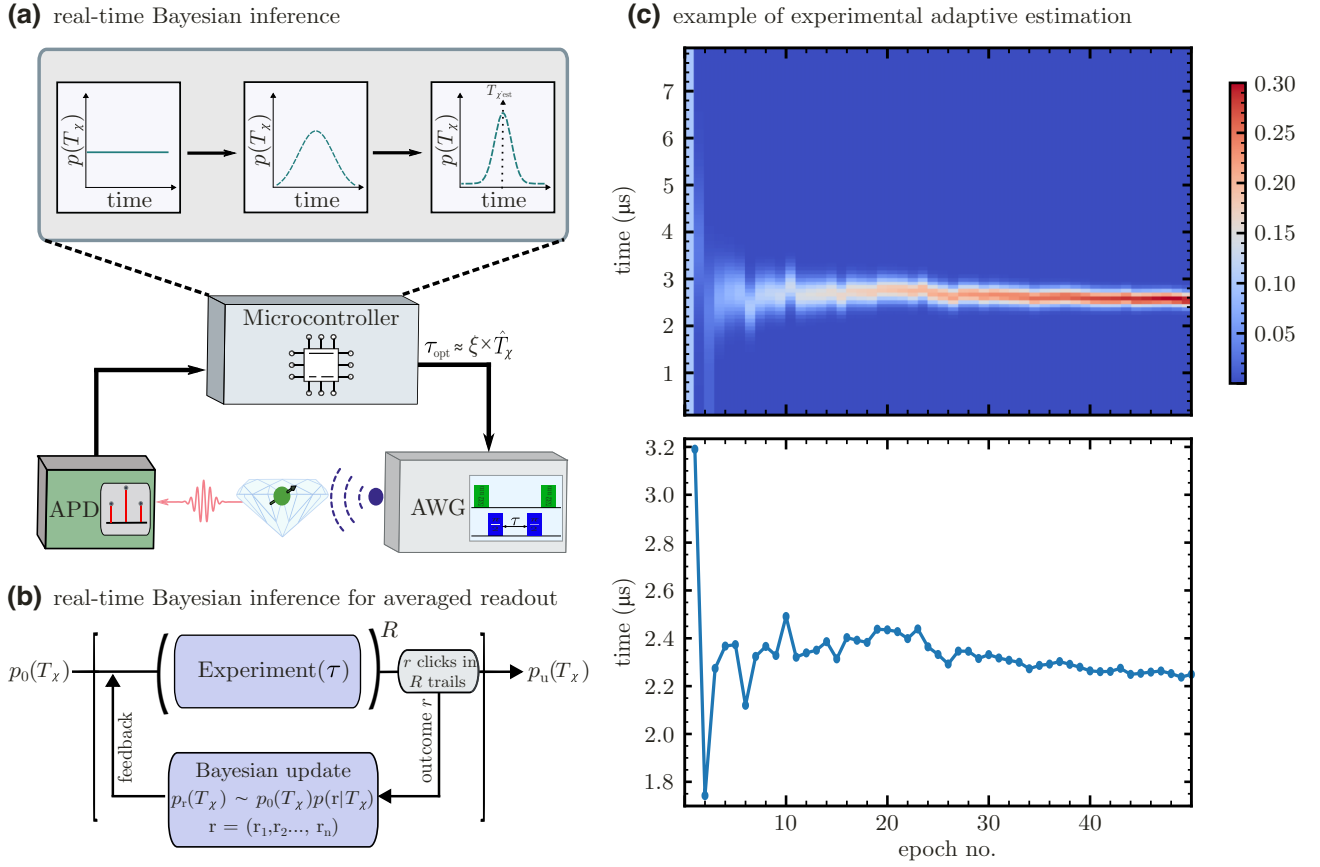


FIG. 1. Real-time adaptive feedback. (a) Schematic of the real-time adaptive protocol demonstrated in this work, using the electronic spin associated with a nitrogen-vacancy (N-V) center in diamond. An arbitrary waveform generator (AWG) is used to generate pulses for manipulation of the spin qubit. The spin state is then optically measured and the detected photon count rate is used by the microcontroller to estimate the value of the decoherence timescale  $T_\chi$  (and the decay exponent  $\beta$ ) using Bayesian inference. The microcontroller also computes the optimal probing time  $\tau$  for the next measurement, passing the value to the AWG, which builds the next estimation sequence accordingly. (b) In our experiment, a single interrogation of the qubit does not provide sufficient information to discriminate qubit state. Hence  $R$  measurements are performed, and the resulting number  $r$  of detected photons are used to update  $p(T_\chi)$  through Bayes' rule and to compute  $\tau$  for the next experiment. (c) Example of an experimental adaptive estimation sequence, with  $p(T_2^*)$  shown for each measurement epoch (here the decay exponent is known *a priori* as  $\beta = 2$ ). The probability  $p(T_2^*)$ , initially uniform, converges towards a narrow peak as more measurement outcomes are accumulated. The experimental adaptively optimized values for  $\tau$  are shown on the bottom plot.

$P(m_n | T_\chi)$  is the likelihood function

$$P(m | T_\chi) = \frac{1 + e^{im\pi} e^{-(\tau/T_\chi)^\beta}}{2}. \quad (4)$$

Note that this likelihood depends on  $\tau$  (which we will adjust later) but this dependency is omitted in  $P(m | T_\chi)$  to simplify the notation. Our approach to adaptive estimation is to derive a simple expression for the optimal parameter settings, that can be computed in real time by an analytical formula without adding much extra computation time to the sensing process.

A conventional approach to updating  $\tau$  adaptively would be to use the information gain as criterion [21,52,55]. However, this involves integrals (with respect to  $T_\chi$ )

requiring numerical evaluation and an associated significant computational overhead. Here, we instead employ an approximation of the Bayesian information matrix (BIM) (a  $1 \times 1$  matrix, in the case of a single parameter) [56], which links to the classical Fisher information [57]. While computing the BIM also requires the computation of an integral (more precisely an expectation) with respect to  $T_\chi$ , this can now be easily approximated as explained in Appendix A.

The Cramér-Rao lower bound (CRLB) of  $T_\chi$  represents the minimum reachable variance for any (unbiased) estimator of  $T_\chi$ , and is inversely proportional to the Fisher information  $F$ . Thus, maximizing  $F$  with respect to the control parameter  $\tau$  is expected to improve our estimate of  $T_\chi$ .

As shown in Appendix A, we examine a Bayesian form of the CRLB, computing the corresponding Fisher information  $F_B$  can be computed as

$$F_B(\tau) \approx \frac{\beta^2(\tau/\hat{T}_\chi)^{2\beta}}{\hat{T}_\chi^2 \left[ e^{(2\tau/\hat{T}_\chi)^\beta} - 1 \right]}, \quad (5)$$

where  $\hat{T}_\chi$  is a point estimate of  $T_\chi$  before each measurement.

While we are unable to maximize  $\hat{F}_B(\tau)$  analytically, approximate solutions exist. We found that the heuristic

$$\tau_{\text{opt}} \approx \xi \times \hat{T}_\chi \quad (6)$$

leads to satisfactory results, where  $\xi$  is a parameter that depends on  $\beta$ . Some numerically computed values for  $\xi$  are listed in Table I, for some common values of  $\beta$ .

The Fisher information  $F$  as a function of the ground-truth value for  $T_\chi = T_2^*$  and the probing time  $\tau$ , from Eq. (A2) is plotted in Fig. 2(a).  $F(\tau, T_2^*)$  is normalized by its maximum with respect to  $\tau$ , for each value of  $T_2^*$ . The plot shows clearly that the maximum of  $F(\tau)$  has a linear dependence on  $T_2^*$ , following Eq. (A2) (shown as the red dashed line).

## B. Multiparameter estimation

In many practical situations, it is worthwhile to learn both the decoherence timescale,  $T_\chi$ , and the noise

TABLE I. Choice of optimal probing time. The optimal probing time  $\tau_{\text{opt}}$  is chosen according to Eq. (A5), with factor  $\xi$ . Numerically computed values of  $\xi$  for different decay exponents  $\beta$  are shown in the table, computed by maximizing either the classical Fisher information  $F$  [see Eq. (A2)] or the rescaled Fisher information  $F_T$  [see Eq. (9)]. No maximum of  $F_T$  can be found for exponential decay, when  $\beta = 1$ .

$\beta$	1	3/2	2	3
$\xi(F)$	0.79	0.86	0.89	0.92
$\xi(F_T)$	...	0.30	0.66	0.85

exponent,  $\beta$ , as the latter provides useful information about the nature of the qubit environment.

Directly extending the approach for a single-parameter estimate is unfruitful as the determinant of the BIM is zero, as reported previously [38]. The intuitive reason for this is that the determination of two parameters from measurements in one single setting creates correlations between the two parameters, resulting in a singular Fisher-information matrix.

We address this issue by using two consecutive measurement results,  $m$  and  $n$ , taken at times  $\tau_0$  and  $\tau_1$ . In this case, a nonzero determinant can be found by maximizing the BIM on  $p(m | T_\chi, \beta) \times p(n | T_\chi, \beta)$ . Assuming binary measurement outcomes, the determinant of this matrix is given by

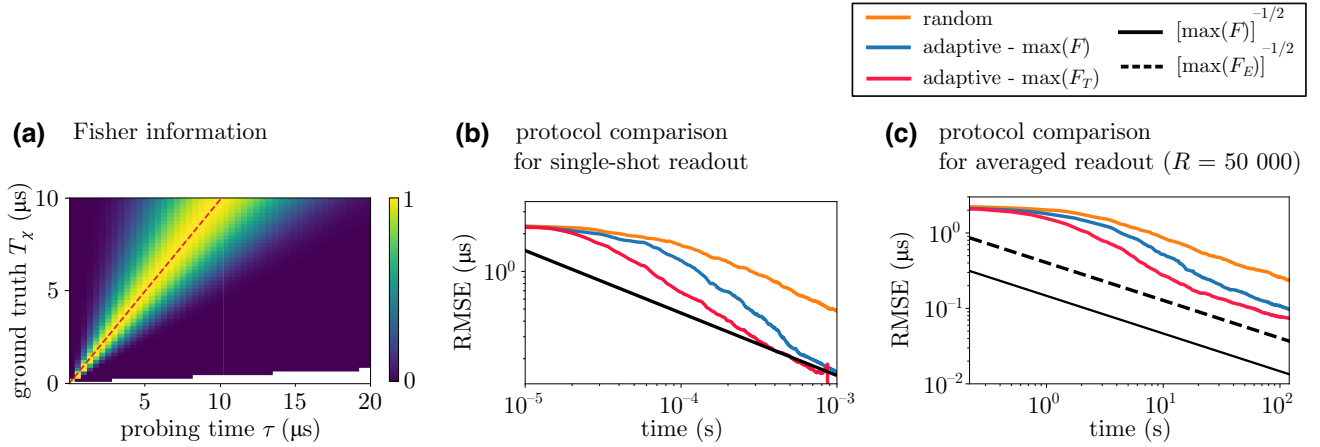


FIG. 2. Numerical simulations for  $T_2^*$  estimation. (a) Fisher information  $F$  as a function of the ground-truth value for  $T_\chi = T_2^*$  and the probing time  $\tau$ , from Eq. (A2). The plot shows that the maximum of  $F(\tau)$  has a linear dependence on  $T_2^*$ , following Eq. (A2) (shown as the red dashed line). (b) Comparison between different strategies to learn  $T_2^*$ , assuming that single-shot qubit readout is available: random choice of  $\tau$  (orange), choice of  $\tau$  obtained maximizing the Fisher information  $F$  in Eq. (A2) (blue), and choice of  $\tau$  obtained maximizing the rescaled Fisher information  $F_T$  in Eq. (9) (red). The x axis shows the total probing time (cumulative over epochs), while the y axis shows the uncertainty, defined as the root-mean-squared error (RMSE) from the ground-truth value. The solid black line corresponds to the theoretical limit set by the Cramér-Rao lower bound (CRLB) for the Fisher information in Eq. (A2). Adaptive protocols outperform the nonadaptive (random) protocol. (c) Simulations for the same protocols as in (b), for the case where single-shot readout is not available [using experimental values  $p_{cl}(0) = 0.0187$  and  $p_{cl}(1) = 0.0148$ , and  $R = 50\,000$ ]. The solid black line corresponds to the limit set by the CRLB for the Fisher information in Eq. (A2). The dashed black line shows the CRLB limit for the Fisher information  $F_E$  in Eq. (10).



$$\det \hat{F}_B = \frac{\beta^2 \left(\frac{\tau_0}{T_\chi}\right)^{2\beta} \left(\frac{\tau_1}{T_\chi}\right)^{2\beta} \left(\log^2 \left(\frac{\tau_0}{T_\chi}\right) - 2 \log \left(\frac{\tau_0}{T_\chi}\right) \log \left(\frac{\tau_1}{T_\chi}\right) + \log^2 \left(\frac{\tau_1}{T_\chi}\right)\right)}{T_\chi^2 \left(-\exp \left[2 \left(\frac{\tau_0}{T_\chi}\right)^\beta\right] - \exp \left[2 \left(\frac{\tau_1}{T_\chi}\right)^\beta\right] + \exp \left[2 \left(\frac{\tau_0}{T_\chi}\right)^\beta + 2 \left(\frac{\tau_1}{T_\chi}\right)^\beta\right] + 1\right)}. \quad (7)$$

We did not find a way of calculating the maxima of this function analytically. Instead, here it is approximated numerically to determine the update heuristic. Starting by numerically estimating the value of  $\tau_1/T_\chi$ , which maximizes Eq. (7) across a range of values of  $\tau_0/T_\chi$  and  $\beta$ , we can fit a piecewise linear approximation for the best choice of  $\tau_1$  that works well for  $1 < \beta < 5$ . This approximation is given by

$$\tau_{1\text{opt}} = \begin{cases} 0.313\tau_0 + 1.04\hat{T}_\chi, & \text{if } \tau_0 < 0.83\hat{T}_\chi \\ 0.7\hat{T}_\chi, & \text{if } 0.83\hat{T}_\chi < \tau_0 < 0.96\hat{T}_\chi \\ 0.109\tau_0 + 0.55\hat{T}_\chi, & \text{if } 0.96\hat{T}_\chi < \tau_0 \end{cases}. \quad (8)$$

### C. Optimizing sensitivity

The approach we introduced in the previous section aims at maximizing the BIM as a proxy for minimizing the uncertainty of  $T_\chi$ . However, this criterion does not account for the overall measurement duration. For example, if two measurements (obtained from two different values of  $\tau$ ) lead to similar BIMs, the shorter one of the two should be more favorable, as our goal is to reach the smallest possible uncertainty in the shortest time. We therefore also consider an alternative approach, which, instead of targeting the minimization of the mean-squared error (MSE), is designed to improve the sensitivity, defined as  $\eta^2 = \text{MSE} \times \tau$  [37]. This leads to another criterion that is the BIM rescaled by the probing time:

$$F_T = \frac{\hat{F}_B(\tau)}{\tau} = \frac{\beta^2 (\tau/\hat{T}_\chi)^{2\beta}}{\tau \hat{T}_\chi^2 \left[ e^{2(\tau/\hat{T}_\chi)^\beta} - 1 \right]}. \quad (9)$$

Again, numerically solving for  $(\tau/\hat{T}_\chi)$ , we can find the optimal value  $\tau_{\text{opt}} = \xi(F_T) \times \hat{T}_\chi$ . The bottom row

in Table I shows different values for the multiplication factor  $\xi$  for different  $\beta$  values. We are unable to find a maximum for the rescaled Fisher information when  $\beta = 1$ .

### D. Theoretical limits

The fundamental performance limit for decoherence timescale learning is given by the CRLB for the Fisher information in Eq. (A2). Simulations for the performance of our proposed algorithm for perfect single-shot readout are presented in Fig. 2(b), comparing three strategies to learn  $T_\chi^*$ : random choice of  $\tau$  (orange curve), choice of  $\tau$  obtained maximizing the Fisher information  $F$  in Eq. (A2) (i.e., minimizing the lower bound on the variance, blue curve) and choice of  $\tau$  obtained maximizing the rescaled Fisher information  $F_T$  in Eq. (9) (i.e., minimizing the ‘‘sensitivity,’’ red curve). The solid black line corresponds to the theoretical limit set by the Cramér-Rao lower bound for the Fisher information expression in Eq. (A2). The curves represent the performance averaged over 500 repetitions of each protocol. The simulations show that both adaptive approaches outperform the non-adaptive (random  $\tau$ ) protocol (orange solid line), with the adaptive protocol maximizing  $F_T$  (red line) performing better than the maximization of  $F$  (blue line). Numerical simulations in Fig. 2(b) confirm that the estimator is asymptotically unbiased, as the RMSE becomes smaller and smaller with increasing sensing time, approaching the fundamental estimation limit set by the Cramer-Rao bound.

To disentangle the effect of suboptimal readout from the performance of the proposed estimation algorithm, we can compute the best performance achievable for a given readout strategy. This can be computed from the Cramér-Rao bound, considering the classical Fisher information  $F_E$  for the experimental likelihood given by Eq. (12):

$$F_E(\tau) = - \frac{\alpha V^2 \beta^2 \left(\frac{\tau}{T_\chi}\right)^{2\beta}}{T_\chi^2 \left( \alpha V^2 + 2\alpha V e^{\left(\frac{\tau}{T_\chi}\right)^\beta} - V e^{\left(\frac{\tau}{T_\chi}\right)^\beta} + \alpha e^{2\left(\frac{\tau}{T_\chi}\right)^\beta} - e^{2\left(\frac{\tau}{T_\chi}\right)^\beta} \right)}. \quad (10)$$

Once again, the ultimate lower bound on the variance obtainable in the limit of many identical repetitions (epochs) is proportional to the inverse  $F_E$ . To find the optimal probing time, given experimental values for  $\alpha$  and  $V$ , we, therefore, minimize  $1/F_E$ , displayed as the solid black line in Figs. 2(b) and 2(c).

### III. EXPERIMENTAL IMPLEMENTATION

#### A. Setup and estimation algorithm

We demonstrate the proposed protocols using the single electron spin associated with a N-V center in diamond [58,59], created by a laser-writing method [60,61]. The N-V electron spin is polarized and measured optically, even at room temperature; it can be further controlled by microwave pulses.

Whilst optical measurement at room temperature is available, it is not possible to obtain binary single-shot readout, unlike at cryogenic temperatures [43]. The absence of a single-shot readout can be circumvented by repeating each measurement sequence  $R$  times. At the  $n$ th iteration, the number  $r_n$  of detected photons is then used to update the distribution of  $T_\chi$  as

$$p(T_\chi | r_{(1:n)}, R) \propto p(r_n | T_\chi, R) p(T_\chi | r_{1:(n-1)}, R). \quad (11)$$

The likelihood is a binomial distribution (as a result of  $R$  independent and identical Bernoulli experiments). The probability of detection of a photon click in a single repetition is given by

$$p_D(\tau, T_\chi, \beta) = \alpha \left( 1 + V e^{-(\tau/T_\chi)^\beta} \right). \quad (12)$$

This probability depends on  $V$  and  $\alpha$ , which can be computed from the experimental probabilities to detect a click when the qubit is in  $|0\rangle$  ( $p_{\text{cl}}(|0\rangle)$ ) or  $|1\rangle$  ( $p_{\text{cl}}(|1\rangle)$ ) as

$$V = \frac{p_{\text{cl}}(|0\rangle) - p_{\text{cl}}(|1\rangle)}{p_{\text{cl}}(|0\rangle) + p_{\text{cl}}(|1\rangle)} \quad (13)$$

and  $\alpha = [p_{\text{cl}}(|0\rangle) + p_{\text{cl}}(|1\rangle)]/2$ .

However, since the probability of detecting a photon in one measurement cycle is very small and  $R$  is large,  $p(r_n | T_\chi, R)$  can be well approximated by a Gaussian distribution [20,37]

$$p(r_i | T_\chi, R) \approx \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(r - R \times p_D(\tau, T_\chi, \beta))^2}{2\sigma^2} \right] \quad (14)$$

with  $\sigma = r(R - r)/R$ .

Note that as the adaptive update rule in Eq. (A5) does not depend on the outcome but only on the current estimate of  $T_\chi$  from  $p(T_\chi | r_{1:n})$ , it can be used in the absence of single-shot readout.

The experimental setup is sketched in Fig. 3(a) (more details in Appendix B). A real-time microcontroller performs the update of the probability distribution of  $T_\chi$  via a particle filter and chooses the optimal probing time  $\tau$ , in a timescale of about 50  $\mu\text{s}$ . The value of  $\tau$  is then passed to an AWG, which, in real time, generates the

#### Input:

$p_0(x)$ : prior probability distribution for  $x = T_\chi$ ;  
 $K$ : number of particles;  
 $N$ : number of epochs;  
 $a_{LW}$ : Liu-West re-sampling parameter;  
 $t_{RS}$ : re-sampling threshold

```

1: procedure ADAPTIVEESTIMATION( $n, p_0, N, a_{LW}, t_{RS}$ )
2:   draw  $\{x_k\}$  from  $p_0(x)$ 
3:    $\{\omega_k\} \leftarrow \{1/K\}$ 
4:   for  $i \in 1..N$  do
5:      $\hat{T}_\chi \leftarrow \sum_j \omega_k \cdot x_k$ 
6:      $\tau \leftarrow \xi \cdot \hat{T}_\chi$ 
7:     for  $j \in 1..R$  do
8:        $m_j \leftarrow \text{EXPERIMENT}(\tau)$ 
9:        $r_i \leftarrow \sum_{j=1}^R m_j$ 
10:       $\{\omega_k\} \leftarrow \{\omega_k \cdot p(r_i | T_\chi, R)\}$ 
11:       $\{\omega_k\} \leftarrow \{\omega_k / (\sum_k \omega_k)\}$ 
12:      if  $1 / \sum \omega_k^2 < n \cdot t_{RS}$  then
13:         $\{x_k\} \leftarrow \text{RESAMPLE}(\{x_k\}, \{\omega_k\}, a_{LW})$ 
14:         $\{\omega_k\} \leftarrow \{1/K\}$ 
15:   return  $\hat{T}_\chi$ 
16:
17: procedure RESAMPLE( $\{x_k\}, \{\omega_k\}, a_{LW}$ )
18:    $\mu \leftarrow \sum_k x_k \cdot \omega_k$ 
19:    $\sigma^2 \leftarrow \sum_k x_k^2 \cdot \omega_k - \mu^2$ 
20:    $\mu' \leftarrow a_{LW} \cdot x_k + (1 - a_{LW}) \cdot \mu$ 
21:    $\{x_k\} \leftarrow \text{NORMAL}(\mu', \sigma^2)$ 
22:   return  $\{x_k\}$ 

```

Algorithm 1. Adaptive estimation algorithm

appropriate spin-manipulation pulse sequence (including laser and microwave pulses). A computational latency of 50  $\mu\text{s}$  is a factor 20 faster than previous real-time quantum sensing experiments at room temperature [24], and much shorter than even the shortest measurement timescales in our experiments (a few milliseconds, for  $\tau \sim 0.1 \mu\text{s}$  and  $R = 10^4$ ).

The estimation algorithm is detailed in Algorithm 1. We first choose a discretization of the distribution of  $T_\chi$ , described by a number  $K$  of particles  $\{x_k\}$  distributed according to the prior probability  $p_0(T_\chi)$  (uniform in our case). We initially set the weights of each particle to  $1/K$ , and  $K = 100$ . At each iteration, we compute the mean  $\hat{T}_\chi$  of the distribution (line 5) and use it to set the next probing time to  $\tau = \xi \times \hat{T}_\chi$  (line 6). We then perform the selected experiment  $R$  times, detecting  $r$  photons. The next step updates the probability distribution  $p(T_\chi)$  according to Bayes' rule (line 11), normalizing the distribution (line 12). At this point, if the distribution features large areas with small weights (described by the condition in line 13), is resampled according to the Liu-West algorithm [62]. We compute the variance  $\sigma^2$  and we then sample new particles (line 25) from a Gaussian distribution with variance  $\sigma^2$  and

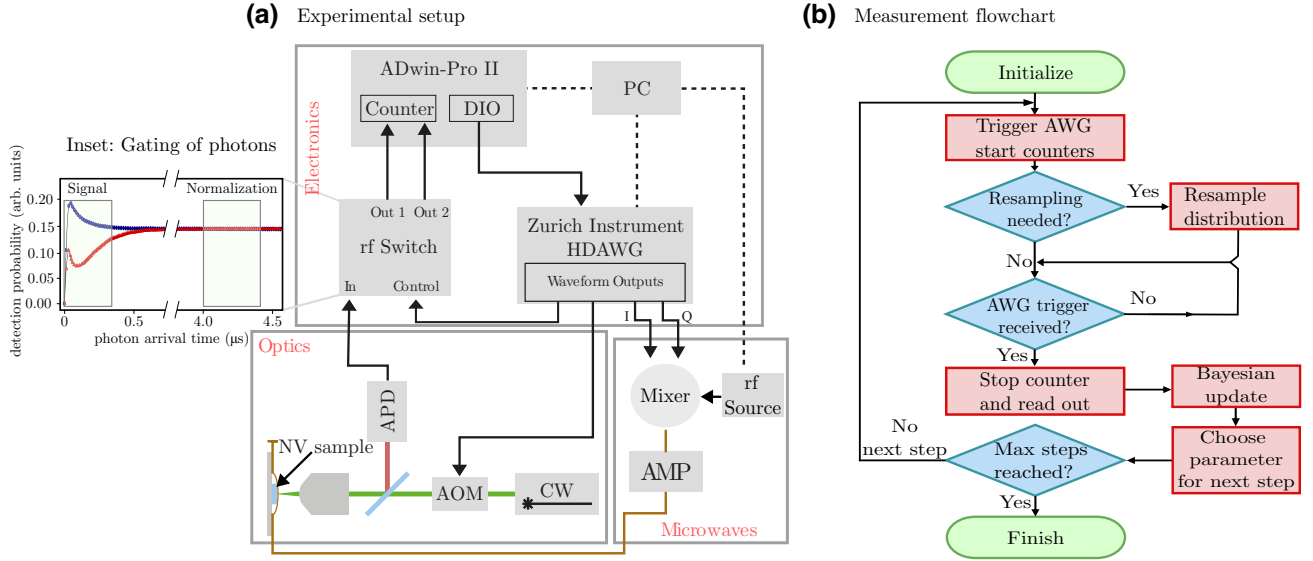


FIG. 3. Experimental setup and measurement flowchart. (a) Schematic of the hardware (optical, microwave, electronics) in the experimental setup with functional connections. A detailed description can be found in Appendix B. Acronyms: CW, continuous wave; AOM, acousto-optic modulator; N-V, nitrogen-vacancy center in diamond; APD, avalanche photodiode; AMP, rf amplifier; AWG, arbitrary waveform generator; PC, computer. The computer is used to set the overall measurement and is then not active during the experiment (managed in hard real time by the AdWin Pro II microcontroller). In the inset, we show the histogram of the photon arrival times when the spin is prepared in  $|0\rangle$  and  $|1\rangle$ . The rf switches select the first shaded range, with a difference in photoluminescence intensity between the two spin states, as signal (directed to counter A) and the second range, with no difference, as background (directed to counter B) to monitor drifts. (b) Flowchart representing the control flow of the adaptive experiments. The control flow of the microcontroller is agnostic to the actual measurement being performed, which is activated by trigger signals. Resampling, if necessary, is performed in parallel to data acquisition (i.e., between the triggering of the AWG with counters starting and the counters stopping) so that no overhead is added to the procedure.

mean:

$$\mu' = a_{LW} \times x_k + (1 - a_{LW}) \times \mu, \quad (15)$$

where  $a_{LW}$  is the Liu-West parameter, which determines how much the next sampling preserves the original  $\{x_k\}$  and how much it reflects the properties (mean  $\mu$ ) of current  $p(T_\chi)$ . Resampling is performed in parallel to data acquisition (see Fig. 3), so that it does not add any additional delay to the overall measurement time.

In the experiments described in Sec. III B we compare the adaptive protocols with a nonadaptive alternative where the probing time  $\tau$  is chosen randomly within an expected range, known *a priori*. The choice of this range affects the performance of the protocols, as adaptive techniques are more effective when there is a large uncertainty in the parameter to be estimated (i.e., for high dynamic range estimation). If the parameter is already known with a good approximation, then settings can already be optimized *a priori*, without the need for real-time adaptation. In the experiments presented in Sec. III B we will select the *a priori* parameter range based on typical ranges of variation of decoherence timescales for N-V centers in diamond at room temperature.

## B. Experimental results

Figure 4 compares the performance of adaptive and nonadaptive protocols to estimate the static dephasing  $T_2^*$  timescale more extensively. We compare three different values of readout repetitions  $R$  ( $R = 10^6$ ,  $R = 10^5$ ,  $R = 10^4$ ). Given our experimentally measured values of  $p_{cl}(|0\rangle) = 0.0186 \pm 0.0017$ , and  $p_{cl}(|1\rangle) = 0.0148 \pm 0.0016$  photons detected per readout for the two qubit states  $\{|0\rangle, |1\rangle\}$ , the three different numbers of readout repetitions  $R$  correspond, respectively, to mean photon numbers  $\langle n \rangle \sim 16,700$ ,  $\langle n \rangle \sim 1670$ , and  $\langle n \rangle \sim 167$ . We implement two forms of nonadaptive protocols, one in which  $\tau$  is chosen at random within the given range 0–8  $\mu\text{s}$  (based on typical experimentally observed ranges), and one in which  $\tau$  is swept uniformly across the range. For a fair comparison, the total number of measurements  $N \times R$  (with  $N$  being the number of epochs) was kept constant for different  $R$ , in order to keep to total measurement time fixed. All curves are averaged over 110 repetitions of the whole estimation sequence in order to obtain the mean performance and the 95% confidence interval.

Figure 4(a) highlights the comparison of measurement uncertainty versus total measurement time for  $R = 10^6$  repetitions per experiment, starting from a flat prior



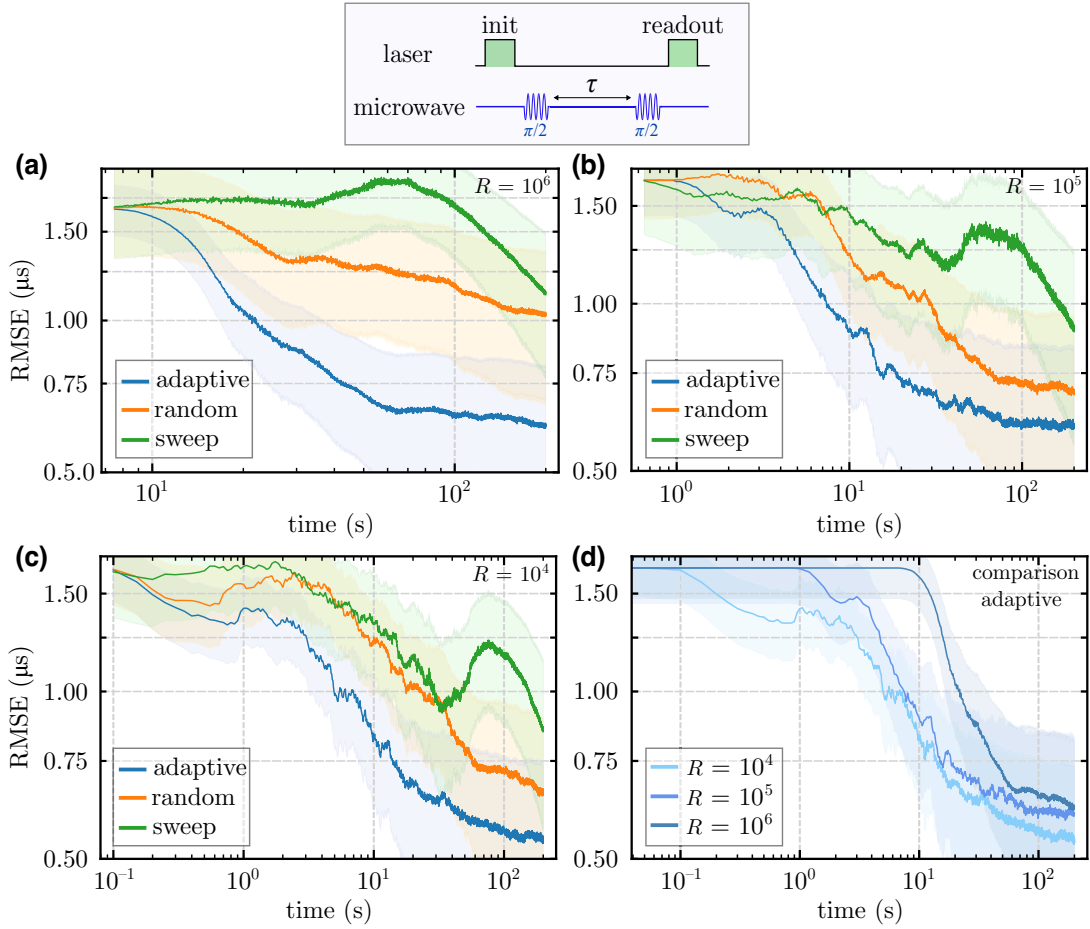


FIG. 4. Experimental comparison between the adaptive and nonadaptive estimation of the dephasing time  $T_2^*$ . The dephasing time  $T_2^*$  is measured by a Ramsey experiment. The spin state, initially polarized by a laser pulse in  $|0\rangle$ , is prepared in an initial superposition state  $(|0\rangle + |1\rangle)/2$  with a microwave  $\pi/2$  pulse. A second  $\pi/2$  pulse, after a delay  $\tau$ , converts the phase information into the population of the  $|0\rangle$  state, which is then readout optically. Our adaptive protocol (blue line) outperforms nonadaptive protocols, either sweeping  $\tau$  on a predetermined range (green line), or randomly picking  $\tau$  (orange line), for a different number of readout repetitions:  $R = 10^6$  (a),  $R = 10^5$  (b), and  $R = 10^4$  (c). In each subplot, we show the RMSE, averaged over 110 experimental protocol repetitions, as a function of the total probing time in seconds. (d) Comparison of the performance of the adaptive protocol for different values of  $R$ , with the same experimental data as in (a)–(c) plotted together for ease of comparison. For all plots the parameter range is  $T_2^* \in [1 \times 10^{-7} \text{ s}, 8 \times 10^{-6} \text{ s}]$ , the ground truth is  $2.5 \mu\text{s}$  and the shaded areas correspond to 95% confidence intervals.

corresponding to the same starting point for all three curves. The evolution of the uncertainty is visibly distinct for the three protocols. The adaptive protocol starts learning about the unknown parameter within just a few epochs, leading to a quick drop in the uncertainty compared to the nonadaptive protocols. For the random choice of delay  $\tau$  the learning is intermediate in performance between the adaptive protocol and the nonadaptive sweep. The sweep protocol performs particularly badly in the first few estimation epochs, as it performs measurements with probing times much shorter than the decoherence rate, resulting in very little information gained. Once it sweeps across probing times closer to the true decoherence rate, uncertainty is reduced much faster. The protocol with random probing times, on the other hand, samples from the whole

expected range uniformly, so that it has a higher probability of retrieving significant information even in the first epochs. The adaptive protocol outperforms both nonadaptive ones, as it quickly learns the optimal probing time to retrieve the most information about the decoherence rate. Comparison between the three experimental curves highlights that, for example, choosing an uncertainty of  $1 \mu\text{s}$ , the adaptive scheme is about 10 times faster than both the nonadaptive schemes.

Figure 4(b) shows the comparison for  $R = 10^5$ , with a similar trend, and the learning achieved by the adaptive protocol is about 4 times faster than the random. The sweep protocol performs a factor of 20 times worse. A similar advantage of the adaptive protocol is also shown for  $R = 10^4$  [see Fig. 4(c)].

In Fig. 4(d), a comparison of the performance of the adaptive scheme for different  $R$  values, shows that  $R = 10^4$  (corresponding to  $\langle n \rangle \sim 167$  photons detected on average) achieves the minimum uncertainty obtained for any measured probing time, highlighting that more frequent adaptive updates are beneficial. Decreasing  $R$  even further, in the Gaussian approximation in Eq. (14), does not improve performance further. Therefore, we fix  $R = 10^4$  for the experiments in the rest of the paper. We also focus only on the version of the nonadaptive protocol with a random choice of  $\tau$ , as it performs better than parameter sweeping.

Results in Fig. 4 show that the sweep protocol performs poorly, due to very little information gain at the start. Therefore, in the subsequent measurements (Figs. 5 and 6) we restricted ourselves to a comparison between the random (nonadaptive) and adaptive schemes. In Fig. 5, we compare the adaptive and nonadaptive protocols for the estimation of  $T_1$  (relaxation) and  $T_2$  (echo decay) timescales. In Fig. 5(top) plot we find that for  $T_1$  estimation the adaptive protocol performs approximately 2 times better than nonadaptive random scheme. This performance is limited by experimental uncertainties, as the long delay time combined with  $R = 10^4$  repetitions leads to greater fluctuations over time as opposed to the other estimation sequences ( $T_2, T_2^*$ ) presented here, where the targeted decoherence timescale in  $\mu\text{s}$  regime. In Fig. 5(bottom) plot we see that the adaptive protocol outperforms the nonadaptive by significant margin. Such a remarkable gain in estimating  $T_2$  of an N-V center in bulk diamond is achieved because the value of  $T_2$  is typically more uncertain *a priori* than the value of  $T_1$ . In both cases, the adaptive strategy greatly outperforms the nonadaptive protocol.

We verify the performance of multiparameter estimation [see Eq. (8)] using Ramsey measurement ( $T_x = T_2^*$ ). The choice of optimal sensing time  $\tau$  is based on using the piecewise linear approximation, discussed in Appendix B. In Fig. 6, we compare the adaptive protocol against random measurement scheme, plotting the estimation error of both  $T_2^*$  and  $\beta$  versus total probing time, demonstrating a clear advantage for the adaptive algorithm. The gain observed in the  $T_2^*$  plot is approximately 10 times for adaptive estimation over the random nonadaptive scheme. While the  $\beta$  estimation plot shows a gain of approximately 20 times for the adaptive scheme. For this experiment the number of particles ( $K$ ) over which the Bayesian update is performed was increased to 500 to better sample the two parameter space.

Figure 7 compares adaptive estimation of the dephasing time  $T_2^*$  when maximizing the BIM (related to  $F$ ) or the BIM rescaled by the probing time (related to  $F_T$ ). As discussed, this results in different multiplication factor  $\xi$  for the adaptive choice of  $\tau$  rule as given in Table I. Figure 7 shows that maximization of  $F_T$  is a factor approximately 2 faster than maximization of  $F$ , as a function of the

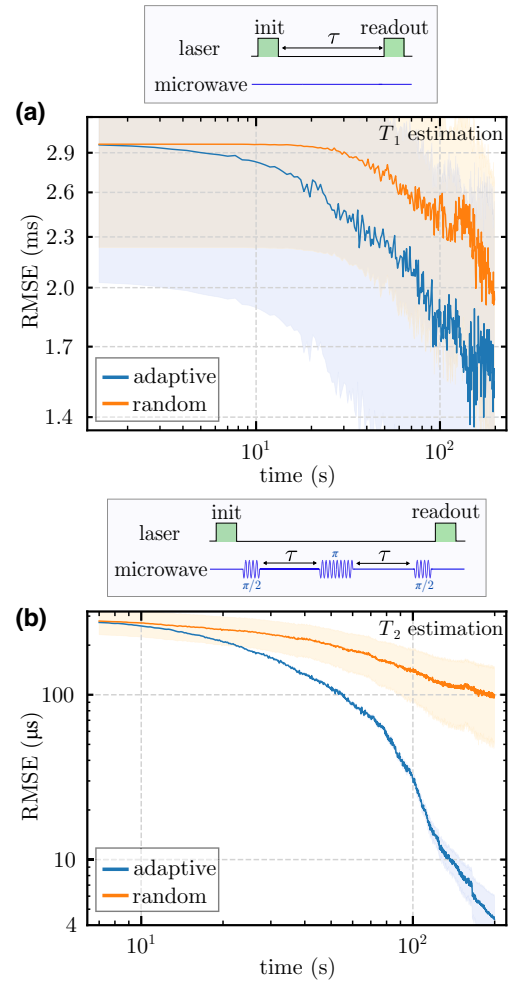


FIG. 5. Experimental comparison of adaptive and nonadaptive estimation for  $T_1$  and  $T_2$ . In both cases the adaptive choice of  $\tau$  is performed maximizing  $F$ , and the number of readout repetitions is  $R = 10^4$ . (a) Standard deviation as a function of probing time in the estimation of spin relaxation  $T_1$  ( $\beta = 1$ ). Each measurement is performed by initializing the electron spin in  $|0\rangle$  and detecting the probability of being in  $|0\rangle$  after a delay  $\tau$ . The parameter range is  $[1 \times 10^{-4} \text{ s}, 8 \times 10^{-3} \text{ s}]$  and the ground truth is 1.45 ms. The standard deviation is averaged over ten repetitions of the protocol. (b) Standard deviation as a function of probing time in the estimation of the Hahn-echo decay time  $T_2$  ( $\beta = 3/2$ ). Each measurement is performed by preparing the qubit in  $(|0\rangle + |1\rangle)/2$  and detecting the overlap with the initial state after a time  $2 \times \tau$ , with a spin flip at time  $\tau$  to rephase the state evolution canceling the effect of static magnetic uncertainty. The parameter range is  $[5 \times 10^{-6} \text{ s}, 1 \times 10^{-3} \text{ s}]$  and the ground truth is 35  $\mu\text{s}$ . The standard deviation is averaged over 40 repetitions of the protocol. In all plots, shaded regions correspond to the 95% confidence interval. The confidence interval is much wider for  $T_1$  estimation as the number of protocol repetitions is quite small (due to the long time required given the ms-scale delays).

total probing time, as expected from Sec. II C. It is worth noting that both protocols shown in Fig. 7 eventually reach the same scaling as the Fisher Information, but with

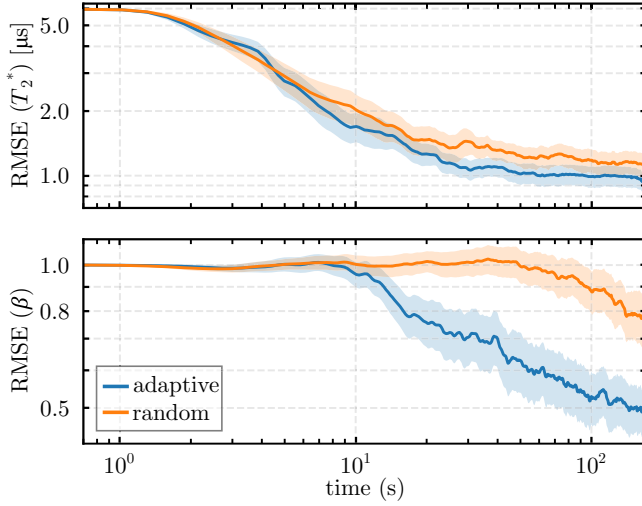


FIG. 6. Simultaneous experimental estimation of  $T_2^*$  (top) and  $\beta$  (bottom). The number of readout repetitions is  $R = 10^5$ . The results are averaged across 140 measurements. The piecewise linear approximation for the adaptive protocol is used. We set our parameter ranges to be  $T_2^* \in [2 \times 10^{-7} \text{ s}, 2 \times 10^{-5} \text{ s}]$ , and  $\beta \in [1, 5]$ . The ground-truth values for our system are  $T_2^* = 4.15 \mu\text{s}$  (a different N- $V$  compared to single-parameter experiments) and  $\beta = 2$ . In all plots, the shaded areas correspond to 95% confidence intervals.

different offsets. This experimental result is consistent with the theoretical prediction from Fig. 2. Similarly, the experimental results shown earlier display strong correspondence with theoretical predictions, providing a robust base for the improved performance of adaptive protocols.

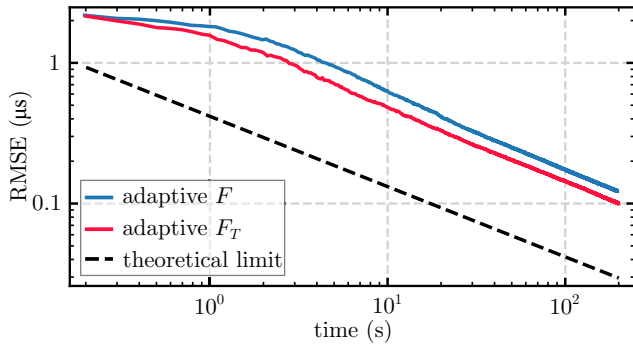


FIG. 7. Experimental comparison between maximizing  $F$  and  $F_T$ . An adaptive choice computed maximizing the Fisher information rescaled by the total probing time [see Eq. (9)] achieves a smaller uncertainty for a given probing time than simply maximizing the Fisher information. The experimental estimations of  $T_2^*$  ( $\beta = 2$ ) are performed with  $R = 10^4$ , with parameter range  $T_2^* \in [1 \times 10^{-7} \text{ s}, 8 \times 10^{-6} \text{ s}]$ , ground truth  $2.5 \mu\text{s}$  and averaged over 35 repetitions of the whole protocol. The dashed black line corresponds to the theoretical limit given by  $F_E$  in Eq. (10). Both experimental curves match the scaling dictated by the CRLB for  $F_E$ .

Generally, while all experimental datasets reproduce the predicted trends, they do not appear to reach the theoretical limits. We believe this is due to residual experimental imperfections or fluctuations not accounted for by the likelihood function we use in the Bayesian inference process, which could be incorporated through data-driven or “gray-box” approaches [63].

#### IV. DISCUSSION AND OUTLOOK

Using an adaptive Bayesian approach based on high-speed electronics, we have compared approaches to estimating decoherence timescales and decay exponents, showing that real-time adaptation of the probing time  $\tau$  significantly outperforms nonadaptive approaches. We have also investigated different adaptation heuristics, demonstrating an advantage in sensitivity optimisation in situations where time is the resource to be minimized.

Real-time adaptation of experimental settings is most valuable when there is a large uncertainty in the parameter to be estimated. In this case, preoptimizing the experiment offline is not trivial, as the optimal settings typically depend on the unknown value of the parameter we are measuring. On the other hand, real-time adaptation is less relevant when the parameter to be estimated varies on a very small range, as in this case the experimental settings can be optimized offline based on the strong available prior information. In this paper, we have chosen reasonable ranges for the decoherence timescales and exponents we estimate, based on our own experience in the field and the variability observed in the literature.

The techniques demonstrated here will directly improve the performance of quantum relaxometry, where measuring the decoherence of a qubit is used to extract useful information from noise in the environment. An example is the use of spins in nanodiamonds to measure the concentration of radicals, exploiting the fact that they induce magnetic noise that shortens the sensor spin  $T_1$  timescale. This has been applied, for example, to probe specific chemical reactions [64] or the concentration of radical oxygen species inside living cells [13,65]. The time required to take a measurement limits the bandwidth over which fast signal variations can be tracked, a very useful parameter when dealing for example with biological processes in living cells.

Our technique can be readily extended to more complex experiments and pulse sequences, for example, double electron-electron resonance (DEER) or electron-nuclear double resonance (ENDOR) where measurements on electronic spins are used to infer the dynamics and decoherence of other electronic or nuclear spin. DEER experiments on single N- $V$  centers close to the diamond surface, for example, reveal changes in decay exponent that elucidate the physics of dark spins on the diamond surface [66].

While our experiments are performed using a single electronic spin associated with an N-V center in diamond, our approach and our analysis are very general and can be readily applied to any qubit system and to different applications. For example, our approach could be used for fast characterization of quantum memories, or multiqubit quantum processors.

In summary, we have shown that real-time adaptive estimation of decoherence, through a simple analytical optimisation heuristic, leads to a considerable speedup compared to nonadaptive schemes. Extending this approach to more complex settings, such as towards the detection of multiple individual nuclear spins and nanoscale magnetic resonance [67–70], holds the promise to deliver even greater advantages.

### ACKNOWLEDGMENTS

We express our gratitude to Mete Atatüre (University of Cambridge) for valuable discussions and appreciate the technical support provided by Andrea Corna and Jelena Trbovic (Zurich Instruments) as well as Heinz Beimert (Jäger Messtechnik GmbH). This work is funded by the Engineering and Physical Sciences Research Council (EP/S000550/1, EP/V053779/1 and through the UK Quantum Technology Hub in Quantum Imaging EP/T00097X/1), the Leverhulme Trust (RPG-2019-388) and the European Commission (QuanTELCO, Grant Agreement No. 862721). We also acknowledge the support provided by a Rank Prize “Return to Research” grant. C.B. and A.F. are jointly supported by the “Making Connections” Weizmann-UK program. C. Bekker is supported by a Royal Academy of Engineering Research Fellowship (No. RF2122-21-129). G.W.M. is supported by the Royal Society (RGF\EA\180311 and UF160400), by the UK EPSRC (EP/V056778/1) and by the UK STFC (ST/W006561/1 and ST/S002227/1). G.W.M and J.S. are jointly supported by the EPSRC Grant No. EP/T001062/1.

### APPENDIX A: DERIVATION OF ADAPTIVE RULES

As discussed in Sec. II A, we optimize the settings for the next measurement by considering an approximation of the Bayesian information matrix (BIM) (a  $1 \times 1$  matrix, in the case of a single parameter) [56], which links to the classical Fisher information [58].

For a binary outcome  $m = 0$  or  $1$ , the classical Fisher information for the decoherence timescale  $T_\chi$  is

$$F = \sum_{m=0,1} \left[ \frac{\partial}{\partial T_\chi} \log p(m | T_\chi) \right]^2 p(m | T_\chi), \quad (\text{A1})$$

where  $p(m | T_\chi)$  is the probability to detect outcome  $m$  given a decoherence timescale  $T_\chi$  described by Eq. (4).

The classical Fisher information  $F$  is then given by

$$F(\tau) = \frac{\beta^2 (\tau/T_\chi)^{2\beta}}{T_\chi^2 \left[ e^{(2\tau/T_\chi)^\beta} - 1 \right]}. \quad (\text{A2})$$

The Fisher information  $F$  is inversely proportional to the CRLB of  $T_\chi$ , which represents the minimum reachable variance for any (unbiased) estimator of  $T_\chi$ . Thus, maximizing  $F$  with respect to  $\tau$  is expected to improve our estimate of  $T_\chi$ . The value of  $\tau$  that maximizes Eq. (A2) (i) needs to be approximated numerically and what is more (ii) depends on  $T_\chi$ , which is unknown [this dependency is not explicit in the notation  $F(\tau)$  but visible in Eq. (A2)]. In this context, where prior information is known about  $T_\chi$  from previous measurements [see Eq. (3)], the Bayesian CRLB [56] is a more adapted criterion as it averages the Fisher information over the high-density region of the prior distribution of  $T_\chi$ . In the single-parameter setting considered here, the BIM is given (see Ref. [56]) by

$$F_B(\tau) = E_{T_\chi} \left[ F(\tau) + \frac{\partial^2 \log(P(T_\chi | m_{1:(n-1)}))}{\partial^2 T_\chi} \right], \quad (\text{A3})$$

where  $E_{T_\chi}[\cdot]$  denotes the statistical expectation with respect to  $P(T_\chi | m_{1:(n-1)})$ . The Bayesian CRLB (BCRLB) is the inverse of  $F_B(\tau)$  and bounds the performance [mean squared error (MSE)] of estimators of  $T_\chi$  [56]. Instead of maximizing the information gain or minimizing the MSE, we use the BIM as a proxy reward function to be maximized when adapting  $\tau$ . First, it is worth noting that the second term in the expectation in Eq. (A3) does not depend on  $\tau$ . Thus, it is sufficient to maximize  $E_{T_\chi}[F(\tau)]$ . To overcome this intractable integral, we use

$$F_B(\tau) \approx \frac{\beta^2 (\tau/\hat{T}_\chi)^{2\beta}}{\hat{T}_\chi^2 \left[ e^{(2\tau/\hat{T}_\chi)^\beta} - 1 \right]} = \hat{F}_B(\tau), \quad (\text{A4})$$

where  $\hat{T}_\chi$  is a point estimate of  $T_\chi$ . This corresponds to replacing, in the integral  $E_{T_\chi}[F(\tau)]$ , the distribution  $P(T_\chi | m_{1:(n-1)})$  by a Dirac  $\delta$  function centered at  $\hat{T}_\chi$ . If the prior distribution  $P(T_\chi | m_{1:(n-1)})$  is concentrated around  $\hat{T}_\chi$ , one option is to choose  $\hat{T}_\chi$  as the most likely value of  $T_\chi$  [15]. Here we used instead the prior mean as it can be more robust when  $P(T_\chi | m_{1:(n-1)})$  has a large variance and/or is skewed. Note that here we do not specify how  $P(T_\chi | m_{1:(n-1)})$  and its moments are computed. This is detailed in Sec. III A.

As discussed in Sec. II A, we find an approximate maximum of Eq. (A4) as

$$\tau_{\text{opt}} \approx \xi \times \hat{T}_\chi, \quad (\text{A5})$$

where  $\xi$  is a parameter that depends on  $\beta$ .



## APPENDIX B: EXPERIMENTAL SETUP

### 1. Sample

The sample consists of an electronic-grade CVD diamond plate (Element Six). Isolated N- $V$  centers with good spin coherence are created by the laser writing of vacancies at an array of sites in the diamond followed by thermal annealing to form N- $V$  centers with background nitrogen impurities.

### 2. Optics

The system is based on a custom-made confocal microscope. The sample, mounted on a custom PCB on a Newport stage, is imaged through an oil immersion objective (Olympus RMS100X-PFO). Photoluminescence is excited by a CW 532-nm diode-pumped solid-state laser (CNI MLL-U-532-200mW), pulsed by an acousto-optic modulator (Isomet 1250C-829A) in double-pass configuration. The laser beam is scanned across the sample by a galvo mirror pair (Thorlabs GVS012/M), using a  $4f$  optical system. The photoluminescence is collected through a dichroic mirror (Semrock LP02-633RU-25) and detected by a fiber-coupled single photon avalanche photodiode (Excelitas SPCM-AQRH-15-FC). The optical setup is mounted inside a styrofoam box and temperature stabilized by a Peltier element controlled by a PID loop (Meerstetter TEC-1092), down to 10 mK rms.

### 3. Magnetic field

A magnetic field of 460 gauss, aligned with the N- $V$  axis, is applied by a permanent SmCo magnet mounted on a 3D motorized translation stage (Standa 8MT173-30DCE2-XYZ). The magnetic field is selected at the N- $V$  center excited-state level anticrossing to polarize the  $^{14}\text{N}$  nuclear spin, so that no hyperfine lines are present in the Ramsey experiments. The magnet is encased in a 3D-printed holder, and thermally stabilised by a Peltier element controlled by a PID loop (Meerstetter TEC-1091), down to 10 mK rms.

### 4. Microwaves

Microwaves are generated by single-sideband modulation of a tone at 1.53 GHz, from a vector source (Rohde-Schwarz SMBV100A), with a low-frequency signal (40 MHz) from a Zurich Instruments HDAWG4 arbitrary waveform generator (2.4 GSa/s, 16 bits vertical resolution, 750 MHz signal bandwidth). The modulated microwave signal is amplified (Amplifier Research 15S1G6, 0.7-6 GHz, 15 W) and directed to a wire across the sample, through a circulator (Aaren Technology 11B-GX017027G-AF), to create an rf magnetic field to drive the ( $m_s = 0 \iff m_s = -1$ ) electron spin resonance.

## 5. Adaptive electronics

The real-time experiment optimization is handled by a hard real-time microcontroller (AdWin Pro II, Jäger Computergesteuerte Messtechnik). Electrical pulses from the photon detectors are sent to a pair of switches (Mini-circuit ZASW-2-50DRA+), that redirect the window A ( $[t_0, t_0 + 270 \text{ ns}]$ ) to counter 1 in the AdWin and the window B ( $[t_0 + 4 \mu\text{s}, t_0 + 4 \mu\text{s} + 400 \text{ ns}]$ ) to counter 2 ( $t_0$  is the start time of the excitation laser pulse). Counts in window A correspond to photon counts that discriminate the spin signal [see Fig. 3(a) inset], since  $m_s = 0$  and  $m_s = \pm 1$  exhibit different photoluminescence intensities. Counts in window B are not spin dependent and are used to monitor the system's stability.

The microcontroller reads the values in the two counters and uses the signal counts in window A to perform a Bayesian update and to select the optimal probing time  $\tau$  for the next measurement. The value of  $\tau$  is passed, as an 8-bit integer to the HDAWG through its digital IO port. For each batch of  $R$  measurements, the HDAWG reads the value from the DIO port and constructs the next pulse sequence accordingly. The HDAWG generates control pulses for the acousto-optic modulator, single-sideband modulation of the microwave signal and to control the switch to redirect electrical pulses from the detector to the microcontroller counters.

### 6. Details of microcontroller (MCU) operation flow

The operation of the MCU is described by the flowchart in Fig. 3(b). At each iteration, the MCU triggers the AWG, which delivers all the control pulse sequences, and starts the counters. If resampling is needed, it is started at this point and carried out in parallel to data acquisition so it does not add any additional overhead. Once the MCU receives a trigger back from the AWG, signaling that all pulse sequences have been executed, it stops and reads the counters, which have accumulated the total count rate for all  $R$  repetitions. The MCU then performs the Bayesian update of  $p(T_\chi)$  and selects the next  $\tau$ , sending the value to the AWG to compile the next qubit control sequence.

The timing of the experiment was benchmarked with respect to the internal clock of the microcontroller, finding  $T_{\mu\text{s}} \sim 0.255 \times n$ , where  $n$  is the number of particles discretising  $T_\chi$ , i.e., about 50  $\mu\text{s}$  for  $n = 200$  particles.

### 7. Details of multiparameter approximate adaptive rule

As we could not find an analytical solution to create an update rule, instead we evaluate the equation numerically across our region of interest and fit an approximation. The results are shown in Fig. 8. We numerically evaluate Eq. (7), modified to include terms to quantify measurement readout occurring over  $R$  repetitions in the absence



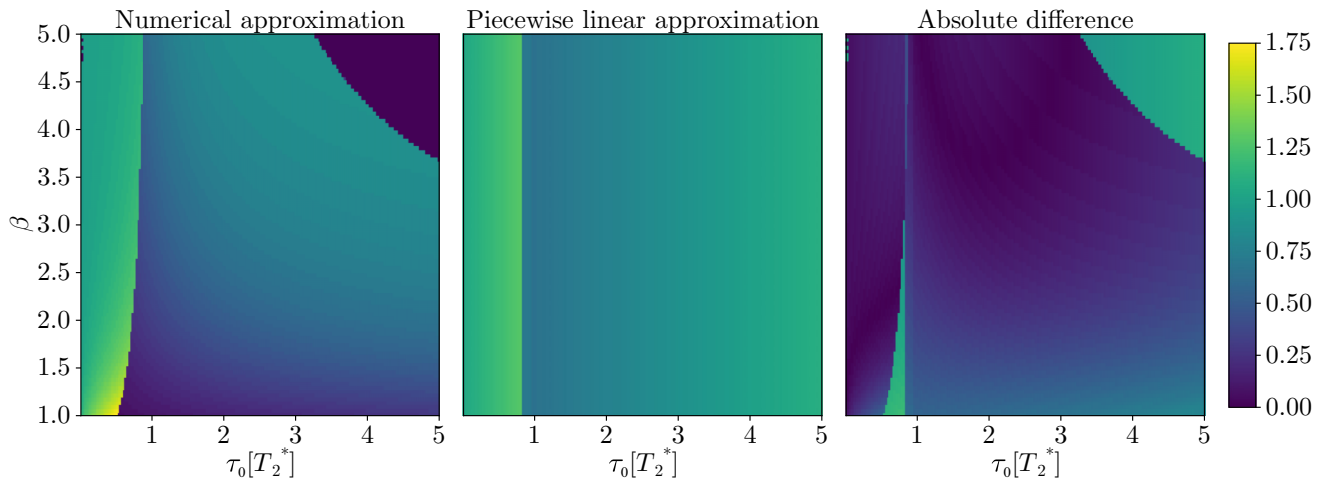


FIG. 8. Piecewise linear approximation accuracy. Plots demonstrate the optimal choice of next measurement time ( $\tau_1$ ) versus the current estimate of coherence decay exponent  $\beta$  and previous measurement time  $\tau_0$ . Both measurement times are normalized to the current estimate of  $T_2^*$ . The accuracy of the numerical approximation (left) is limited by the 64-bit resolution of the floating point used in the calculation, resulting in a region in the upper right of the numerical approximation plot that was manually set to zero. The piecewise linear approximation given by Eq. (8) is plotted in the center, and the absolute difference between these two approximations is given in the right-hand-side plot.

of single-shot readout. We aim to find the simplest approximation to replicate its result. Observation of the steep jump in value just below  $\tau_0 = 1$  suggested the use of a piecewise polynomial fit in  $\tau_0$  and  $\beta$ . Using the polyfit function of the numpy package we find optimal parameters for linear fits in two separate regions. In between these regions, a heuristic choice is made to pick a measurement time  $0.7 \times T_2^*$ . Simulation of experiments using this piecewise linear approximation rule showed similar performance to the use of the full numerical approximation.

- 
- [1] D. Suter and G. A. Álvarez, Colloquium: Protecting quantum information against environmental noise, *Rev. Mod. Phys.* **88**, 041001 (2016).
- [2] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, *Science* **362**, eaam9288 (2018).
- [3] M. Pompili, S. L. N. Hermans, S. Baier, H. K. C. Beukers, P. C. Humphreys, R. N. Schouten, R. F. L. Vermeulen, M. J. Tiggelman, L. dos Santos Martins, B. Dirkse, *et al.*, Realization of a multinode quantum network of remote solid-state qubits, *Science* **372**, 259 (2021).
- [4] C. E. Bradley, J. Randall, M. H. Aboeib, R. C. Berrevoets, M. J. Degen, M. A. Bakker, M. Markham, D. J. Twitchen, and T. H. Taminiau, A ten-qubit solid-state spin register with quantum memory up to one minute, *Phys. Rev. X* **9**, 031045 (2019).
- [5] M. Zhong, M. P. Hedges, R. L. Ahlefeldt, J. G. Bartholomew, S. E. Beavan, S. M. Wittig, J. J. Longdell, and M. J. Sellars, Optically addressable nuclear spins in a solid with a six-hour coherence time, *Nature* **517**, 177 (2015).
- [6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. GSL Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [7] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.*, Strong quantum computational advantage using a superconducting quantum processor, *Phys. Rev. Lett.* **127**, 180501 (2021).
- [8] A. Crippa, R. Ezzouch, A. Aprá, A. Amisse, R. Laviéville, L. Hutin, B. Bertrand, M. Vinet, Matias Urdampilleta, Tristan Meunier, *et al.*, Gate-reflectometry dispersive readout and coherent control of a spin qubit in silicon, *Nat. Commun.* **10**, 1 (2019).
- [9] M. Vinet, The path to scalable quantum computing with silicon spin qubits, *Nat. Nanotechnol.* **16**, 1296 (2021).
- [10] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, *Rev. Mod. Phys.* **89**, 035002 (2017).
- [11] D. Schmid-Lorch, T. Häberle, F. Reinhard, A. Zappe, M. Slota, L. Bogani, A. Finkler, and J. Wrachtrup, Relaxometry and dephasing imaging of superparamagnetic magnetite nanoparticles using a single qubit, *Nanoletters* **15**, 4942 (2015).
- [12] A. Finco, A. Haykal, R. Tanos, F. Fabre, S. Chouaieb, W. Akhtar, I. Robert-Philip, W. Legrand, F. Ajejas, K. Bouzehouane, *et al.*, Imaging non-collinear antiferromagnetic textures via single spin relaxometry, *Nat. Commun.* **12**, 1 (2021).
- [13] L. Nie, A. C. Nusantara, V. G. Damle, R. Sharmin, E. P. P. Evans, S. R. Hemelaar, K. J. Van der Laan, R. Li, F. P. Perona Martinez, T. Vedelaar, *et al.*, Quantum monitoring of cellular metabolic activities in single mitochondria, *Sci. Adv.* **7**, eabf0573 (2021).
- [14] T. Zhang, G. Pramanik, K. Zhang, M. Gulka, L. Wang, J. Jing, F. Xu, Z. Li, Q. Wei, P. Cigler, *et al.*, Toward

- quantitative bio-sensing with nitrogen-vacancy center in diamond, *ACS Sens.* **6**, 2077 (2021).
- [15] Y.-H. Zhang and W. Yang, Improving spin-based noise sensing by adaptive measurements, *New J. Phys.* **20**, 093011 (2018).
- [16] V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Bianchi, F. Marquardt, L. Pezzè, and C. Bonato, Learning quantum systems, *Nat. Rev. Phys.* **5**, 141 (2023).
- [17] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, *Nature* **450**, 393 (2007).
- [18] R. S. Said, D. W. Berry, and J. Twamley, Nanoscale magnetometry using a single-spin system in diamond, *Phys. Rev. B* **83**, 125410 (2011).
- [19] 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).
- [20] I. Zohar, B. Haylock, Y. Romach, M. J. Arshad, N. Halay, N. Drucker, R. Stöhr, A. Denisenko, Y. Cohen, C. Bonato, and et al., Real-time frequency estimation of a qubit without single-shot-readout, *Quantum Sci. Technol.* **8**, 035017 (2023).
- [21] R. D. McMichael, S. Dushenko, and S. M. Blakley, Sequential Bayesian experiment design for adaptive Ramsey sequence measurements, *J. Appl. Phys.* **130**, 144401 (2021).
- [22] R. D. McMichael and S. M. Blakley, Simplified algorithms for adaptive experiment design in parameter estimation, *Phys. Rev. Appl.* **18**, 054001 (2022).
- [23] R. Santagati, A. A. Gentile, S. Knauer, S. Schmitt, S. Paesani, C. Granade, N. Wiebe, C. Osterkamp, L. P. McGuinness, J. Wang, *et al.*, Magnetic-field learning using a single electronic spin in diamond with one-photon readout at room temperature, *Phys. Rev. X* **9**, 021019 (2019).
- [24] T. Joas, S. Schmitt, R. Santagati, A. A. Gentile, C. Bonato, A. Laing, L. P. McGuinness, and F. Jelezko, Online adaptive quantum characterization of a nuclear spin, *Quantum Inf.* **7**, 1 (2021).
- [25] M. Caouette-Mansour, A. Solyom, B. Ruffolo, R. D. McMichael, J. Sankey, and L. Childress, Robust spin relaxometry with fast adaptive Bayesian estimation, *Phys. Rev. Appl.* **17**, 064031 (2022).
- [26] 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).
- [27] C. Bonato and D. W. Berry, Adaptive tracking of a time-varying field with a quantum sensor, *Phys. Rev. A* **95**, 052348 (2017).
- [28] K. Craigie, E. M. Gauger, Y. Altmann, and C. Bonato, Resource-efficient adaptive Bayesian tracking of magnetic fields with a quantum sensor, *J. Phys. Condens. Matter* **33**, 195801 (2021).
- [29] Jacob Benestad, Jan A. Krzywda, Evert van Nieuwenburg, and Jeroen Danon, Efficient adaptive Bayesian estimation of a slowly fluctuating Overhauser field gradient, *ArXiv:2309.15014* (2023).
- [30] M. Valeri, V. Cimini, S. Piacentini, F. Ceccarelli, E. Polino, F. Hoch, G. Bizzarri, G. Corrielli, N. Spagnolo, R. Osellame, *et al.*, Experimental multiparameter quantum metrology in adaptive regime, *Phys. Rev. Res.* **5**, 013138 (2023).
- [31] A. Lumino, E. Polino, A. S. Rab, G. Milani, N. Spagnolo, N. Wiebe, and F. Sciarrino, Experimental phase estimation enhanced by machine learning, *Phys. Rev. Appl.* **10**, 044033 (2018).
- [32] V. Cimini, M. Valeri, E. Polino, S. Piacentini, F. Ceccarelli, G. Corrielli, N. Spagnolo, R. Osellame, and F. Sciarrino, Deep reinforcement learning for quantum multiparameter estimation, *Adv. Photonics* **5**, 016005 (2023).
- [33] M. S. Blok, C. Bonato, M. L. Markham, D. J. Twitchen, V. V. Dobrovitski, and R. Hanson, Manipulating a qubit through the backaction of sequential partial measurements and real-time feedback, *Nat. Phys.* **10**, 189 (2014).
- [34] R. Vijay, C. Macklin, D. H. Slichter, S. J. Weber, K. W. Murch, R. Naik, A. N. Korotkov, and I. Siddiqi, Stabilizing Rabi oscillations in a superconducting qubit using quantum feedback, *Nature* **490**, 77 (2012).
- [35] E. Scerri, E. M. Gauger, and C. Bonato, Extending qubit coherence by adaptive quantum environment learning, *New J. Phys.* **22**, 035002 (2020).
- [36] M. D. Shulman, S. P. Harvey, J. M. Nichol, S. D. Bartlett, A. C. Doherty, V. Umansky, and A. Yacoby, Suppressing qubit dephasing using real-time Hamiltonian estimation, *Nat. Commun.* **5**, 5156 (2014).
- [37] 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).
- [38] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, Robust online Hamiltonian learning, *New J. Phys.* **14**, 103013 (2012).
- [39] P. Cappellaro, Spin-bath narrowing with adaptive parameter estimation, *Phys. Rev. A* **85**, 030301 (2012).
- [40] A. H. Myerson, D. J. Szwer, S. C. Webster, D. T. C. Allcock, M. J. Curtis, G. Imreh, J. A. Sherman, D. N. Stacey, A. M. Steane, and D. M. Lucas, High-fidelity readout of trapped-ion qubits, *Phys. Rev. Lett.* **100**, 200502 (2008).
- [41] F. Mallet, F. R. Ong, A. Palacios-Laloy, F. Nguyen, P. Bertet, D. Vion, and D. Esteve, Single-shot qubit readout in circuit quantum electrodynamics, *Nat. Phys.* **5**, 791 (2009).
- [42] R. Hanson, L. H. Willems van Beveren, I. T. Vink, J. M. Elzerman, W. J. M. Naber, F. H. L. Koppens, L. P. Kouwenhoven, and L. M. K. Vandersypen, Single-shot readout of electron spin states in a quantum dot using spin-dependent tunnel rates, *Phys. Rev. Lett.* **94**, 196802 (2005).
- [43] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. A. Alkemade, and R. Hanson, High-fidelity projective readout of a solid-state spin quantum register, *Nature* **477**, 574 (2011).
- [44] C. P. Anderson, E. O. Glen, C. Zeledon, A. Bourassa, Y. Jin, Y. Zhu, C. Vorwerk, A. L. Crook, H. Abe, J. Ul-Hassan, T. Ohshima, N. T. Son, G. Galli, and D. D. Awschalom, Five-second coherence of a single spin with single-shot readout in silicon carbide, *Sci. Adv.* **8**, eabm5912 (2022)..
- [45] S. Chen, M. Raha, C. M. Phenicie, S. Ourari, and J. D. Thompson, Parallel single-shot measurement and

- coherent control of solid-state spins below the diffraction limit, *Science* **370**, 592 (2020)..
- [46] Ł. Cywiński, R. M. Lutchyn, C. P. Nave, and S. DasSarma, How to enhance dephasing time in superconducting qubits, *Phys. Rev. B* **77**, 174509 (2008).
- [47] L. T. Hall, J. H. Cole, and L. C. L. Hollenberg, Analytic solutions to the central-spin problem for nitrogen-vacancy centers in diamond, *Phys. Rev. B* **90**, 075201 (2014).
- [48] V. V. Dobrovitski, A. E. Feiguin, D. D. Awschalom, and R. Hanson, Decoherence dynamics of a single spin versus spin ensemble, *Phys. Rev. B* **77**, 245212 (2008).
- [49] R. Hanson, V. V. Dobrovitski, A. E. Feiguin, O. Gywat, and D. D. Awschalom, Coherent dynamics of a single spin interacting with an adjustable spin bath, *Science* **320**, 352 (2008).
- [50] E. Bauch, S. Singh, J. Lee, C. A. Hart, J. M. Schloss, M. J. Turner, J. F. Barry, L. M. Pham, N. Bar-Gill, S. F. Yelin, *et al.*, Decoherence of ensembles of nitrogen-vacancy centers in diamond, *Phys. Rev. B* **102**, 134210 (2020).
- [51] Y. Li, L. Pezzè, M. Gessner, Z. Ren, W. Li, and A. Smerzi, Frequentist and Bayesian quantum phase estimation, *Entropy* **20**, Article number 628 (2018).
- [52] D. T. Lennon, H. Moon, L. C. Camenzind, L. Yu, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, E. A. Laird, and N. Ares, Efficiently measuring a quantum device using machine learning, *Quantum Inf.* **5**, 1 (2019).
- [53] J. D. Teske, S. S. Humpohl, R. Otten, P. Bethke, P. Cerfontaine, J. Dedden, A. Ludwig, A. D. Wieck, and H. Bluhm, A machine learning approach for automated fine-tuning of semiconductor spin qubits, *Appl. Phys. Lett.* **114**, 133102 (2019).
- [54] A. A. Gentile, B. Flynn, S. Knauer, N. Wiebe, S. Paesani, C. E. Granade, J. G. Rarity, R. Santagati, and A. Laing, Learning models of quantum systems from experiments, *Nat. Phys.* **17**, 837 (2021).
- [55] D. V. Lindley, On a measure of the information provided by an experiment, *Ann. Math. Stat.* **27**, 986 (1956).
- [56] H. L. Van Trees and K. L. Bell, Bayesian bounds for parameter estimation and nonlinear filtering/tracking, *AMC* **10**, 10 (2007). <https://ieeexplore.ieee.org/book/5263120>.
- [57] S. M. Kay, in *Fundamentals of statistical signal processing: estimation theory* (Prentice-Hall, Inc., NJ, United States, 1993), p. 595. <https://dl.acm.org/doi/abs/10.5555/151045>.
- [58] R. Schirhagl, K. Chang, M. Loretz, and C. L. Degen, Nitrogen-vacancy centers in diamond: Nanoscale sensors for physics and biology, *Annu. Rev. Phys. Chem.* **65**, 83 (2014).
- [59] J. F. Barry, J. M. Schloss, E. Bauch, M. J. Turner, C. A. Hart, L. M. Pham, and R. L. Walsworth, Sensitivity optimization for NV-diamond magnetometry, *Rev. Mod. Phys.* **92**, 015004 (2020).
- [60] Y.-C. Chen, B. Griffiths, L. Weng, S. S. Nicley, S. N. Ishmael, Y. Lekhai, S. Johnson, C. J. Stephen, B. L. Green, G. W. Morley, *et al.*, Laser writing of individual nitrogen-vacancy defects in diamond with near-unity yield, *Optica* **6**, 662 (2019).
- [61] C. J. Stephen, B. L. Green, Y. N. D. Lekhai, L. Weng, P. Hill, S. Johnson, A. C. Frangeskou, P. L. Diggle, Y. C. Chen, M. J. Strain, E. Gu, M. E. Newton, J. M. Smith, P. S. Salter, and G. W. Morley, Deep three-dimensional solid-state qubit arrays with long-lived spin coherence, *Phys. Rev. Appl.* **12**, 064005 (2019).
- [62] J. Liu and M. West, in *Sequential Monte Carlo methods in practice* (Springer New York, New York, NY, 2001), p. 197. <https://scholars.duke.edu/publication/1024706>.
- [63] A. Youssry, Y. Yang, R. J. Chapman, B. Haylock, F. Lenzini, M. Lobino, and A. Peruzzo, Experimental graybox quantum system identification and control, *npj Quantum Inf.* **10**, 9 (2024).
- [64] T. Rendler, J. Neburkova, O. Zemek, J. Kotek, A. Zappe, Z. Chu, P. Cigler, and J. Wrachtrup, Optical imaging of localized chemical events using programmable diamond quantum nanosensors, *Nat. Commun.* **8**, 1 (2017).
- [65] L. Nie, A. Nusantara, V. G. Damle, M. V. Baranov, M. Chipaux, C. Reyes-San-Martin, T. Hamoh, C. P. Epperala, M. Guricova, P. Cigler, *et al.*, Quantum sensing of free radicals in primary human dendritic cells, *Nano Lett.* **22**, 1818 (2021).
- [66] B. L. Dwyer, L. V. H. Rodgers, E. K. Urbach, D. Bluvstein, S. Sangtawesin, H. Zhou, Y. Nassab, M. Fitzpatrick, Z. Yuan, K. De Greve, E. L. Peterson, H. Knowles, T. Sumarac, J.-P. Chou, A. Gali, V. V. Dobrovitski, M. D. Lukin, and N. P. de Leon, Probing spin dynamics on diamond surfaces using a single quantum sensor, *PRX Quantum* **3**, 040328 (2022).
- [67] J. Zopes, K. S. Cujia, K. S. Sasaki, J. M. Boss, K. M. Itoh, and C. L. Degen, Three-dimensional localization spectroscopy of individual nuclear spins with sub-Angstrom resolution, *Nat. Commun.* **9**, 4678 (2018).
- [68] M. H. Abobeih, J. Randall, C. E. Bradley, H. P. Bartling, M. A. Bakker, M. J. Degen, M. Markham, D. J. Twitchen, and T. H. Taminiau, Atomic-scale imaging of a 27-nuclear-spin cluster using a quantum sensor, *Nature* **576**, 411 (2019).
- [69] K. Jung, M. H. Abobeih, J. Yun, G. Kim, H. Oh, A. Henry, T. H. Taminiau, and D. Kim, Deep learning enhanced individual nuclear-spin detection, *Quantum Inf.* **7**, 41 (2021).
- [70] Raffi Budakian, Amit Finkler, Alexander Eichler, Martino Poggio, Christian L. Degen, Sahand Tabatabaei, Inhee Lee, P. Chris Hammel, Eugene S. Polzik, Tim H. Taminiau, *et al.*, Roadmap on nanoscale magnetic resonance imaging, *ArXiv:2312.08841* (2023).