Improving spin-based noise sensing by adaptive measurements

Yi-Hao Zhang and Wen Yang
Beijing Computational Science Research Center, Beijing 100193, China

Localized spins in the solid state are attracting widespread attention as highly sensitive quantum sensors with nanoscale spatial resolution and fascinating applications. Recently, adaptive measurements were used to improve the dynamic range for spin-based sensing of deterministic Hamiltonian parameters. Here we explore a very different direction – spin-based adaptive sensing of random noises. First, we identify distinguishing features for the sensing of magnetic noises compared with the estimation of deterministic magnetic fields, such as the different dependences on the spin decoherence, the different optimal measurement schemes, the absence of the modulo-2π phase ambiguity, and the crucial role of adaptive measurement. Second, we perform numerical simulations that demonstrate significant speed up of the characterization of the spin decoherence time via adaptive measurements. This paves the way towards adaptive noise sensing and coherence protection.

I. INTRODUCTION

Localized electronic spins in the solid state, such as nitrogen-vacancy centers in diamond [1], phosphorus donors [2], silicon vacancy in SiC [3], and single rare-earth ion in yttrium aluminium garnet [4], are attracting widespread attention as highly sensitive quantum sensors [5] with nanoscale spatial resolution [6–11] and fascinating applications in condensed matter physics, materials science, and biology. The coherent Larmor precession of the spin reveals deterministic magnetic signals [12–16], while the decoherence of the spin reveals random magnetic noises [17–20] and other quantum objects [21–28]. By tracking the noises back to the environment, the localized spin can further reveal the structure and many-body physics of the environments, such as quantum criticality [29, 30] and partition functions in the complex plane [31–35] and the quantum work spectrum [36–38]. In these developments, the key challenge is to improve the sensing precision. For this purpose, dynamical decoupling techniques – originally developed for protecting qubits from decoherence – have been adapted for sensing alternating signals [29, 40], noises [41–45], and other quantum objects [21–28, 46–52]. Other techniques include rotating-frame magnetometry [53–55], Floquet spectroscopy [56], two-dimensional spectroscopy [57, 58], correlative measurements [59], axilary quantum memory [60], and compressive sensing [61–62].

Recently, there were growing interest in using adaptive measurements to mitigate the modulo-2π phase ambiguity and hence improve the dynamic range of spin-based quantum sensing [63–68]. However, previous works focus on deterministic Hamiltonian parameters that drive the unitary evolution of the spin quantum sensor, leaving a large, important family of tasks unexplored – the spin-based quantum sensing of random noises that drive the non-unitary decoherence of the spin. It is important to identify the distinctions of spin-based noise sensing compared with the spin-based Hamiltonian parameter estimation, and further provide feasible methods to improve the key figure of merit – the sensing precision.

In this work, we explore theoretically the role of adaptive measurement in spin-based sensing of magnetic noises. First, our general analysis identifies a series of distinguishing features for sensing a random magnetic field (i.e., magnetic noises) compared with the estimation of a deterministic magnetic field (which is a paradigmatic Hamiltonian parameter), including the different dependences on the spin decoherence, the different optimal measurement schemes, and the absence of the modulo-2π phase ambiguity. Moreover, optimizing noise sensing requires knowledge about the unknown noises to be estimated, so adaptive measurements are crucial for improving the sensing precision. By contrast, in the estimation of deterministic magnetic fields, adaptive measurements are usually alternatives to non-adaptive schemes for mitigating the modulo-2π phase ambiguity and hence improving the dynamic range [64, 65] and non-adaptive measurements can even outperform adaptive ones in some cases [64]. Second, we perform numerical simulations and demonstrate that using adaptive measurements can speed up significantly the estimation of the spin decoherence time. These results pave the way towards spin-based adaptive sensing of noises. Since rapid characterization of decoherence allows us to design efficient schemes to suppress the decoherence, these results are also relevant to quantum computation.

This rest of this paper is organized as follows. In Sec. II, we outline the basic steps of a general adaptive measurement, leaving a detailed introduction to every step in Appendices A-E. In Sec. III, we analyze general spin-based sensing of magnetic noises and identify its distinguishing features. In Sec. IV, we perform numerical simulations for the adaptive estimation of the spin decoherence time. In Sec. IV, we draw the conclusion.

II. ADAPTIVE QUANTUM PARAMETER ESTIMATION

A general parameter estimation protocol using a quantum system to estimate an unknown, real parameter ζ consists of three steps (Fig. 1):

1. The quantum system is prepared into certain (usually non-classical) initial state \( \rho_0 \) and then undergoes certain \( \zeta \)-dependent evolution into a final state \( \rho_\zeta \). This step encodes the information about \( \zeta \) into the final state \( \rho_\zeta \) of the quantum system. The information contained in \( \rho_\zeta \) is quantified by the quantum Fisher information (QFI) \( F(\zeta) \).
2. The quantum system undergoes a measurement, which produces an outcome according to certain probability distribution. In this step, the quantum Fisher information $F(\zeta)$ contained in $\hat{\rho}_c$ is transferred into the classical information in the measurement outcome. The information contained in each outcome is quantified by the classical Fisher information (CFI) $F(\zeta)$, which obeys

$$F(\zeta) \leq \mathcal{F}(\zeta). \quad (1)$$

3. Steps 1-2 are repeated $N$ times and the $N$ outcomes are processed to yield an estimator $\hat{\zeta}_{est}$ to the unknown parameter $\zeta$. In this step, the total CFI $NF(\zeta)$ contained in the $N$ outcomes is converted to the estimation precision, as quantified by the statistical error of the estimator:

$$\delta \zeta \equiv \sqrt{\langle (\zeta_{est} - \zeta)^2 \rangle}, \quad (2)$$

where $\langle \cdots \rangle$ denotes the average over a lot of estimators obtained by repeating steps 1-3 many times. For unbiased estimators obeying $\langle \zeta_{est} \rangle = \zeta$, the precision $\delta \zeta$ is fundamentally limited by the inequality

$$\delta \zeta \geq \frac{1}{\sqrt{NF(\zeta)}} \Leftrightarrow (\delta \zeta)^2 \leq NF(\zeta), \quad (3)$$

known as the Cramér-Rao bound [69, 70].

For optimal performance, it is necessary to optimize each step of the above initialization-evolution-measurement cycle. In step 1, the initial state $\hat{\rho}_m$ and the evolution process should be optimized to maximize $F(\zeta)$. In step 2, appropriate measurements should be designed to convert all the QFI contained in $\hat{\rho}_c$ into the CFI contained in the measurement outcome, so that $F(\zeta)$ attains its maximum value $\mathcal{F}(\zeta)$ allowed by Eq. (1). In step 3, optimal unbiased estimators should be used to convert all the CFI $NF(\zeta)$ contained in the $N$ outcomes into the useful information $(\delta \zeta)^{-2}$ contained in the estimator. For example, for large $N$, the Bayesian estimator or the maximum likelihood estimator are unbiased and can saturate the Cramér-Rao bound Eq. (3).

In step 1 and step 2, $\mathcal{F}(\zeta)$ and $F(\zeta)$ may depend on $\zeta$, so the optimization for maximal $\mathcal{F}(\zeta)$ and $F(\zeta)$ requires knowledge about the true value (denoted by $\zeta_{true}$) of the unknown parameter $\zeta$. A possible solution is adaptive measurements [71, 72], i.e., using the measurement outcomes of previous initialization-evolution-measurement cycles to refine our knowledge about $\zeta$ and then use this knowledge to optimize the next cycle. In step 2-3, the probability distribution of the measurement outcome as a function of the unknown parameter $\zeta$ may be periodic, making it impossible to identify a unique estimator. This ambiguity problem is commonly encountered in estimating deterministic Hamiltonian parameters and can be mitigated by using either non-adaptive or adaptive measurements [64]. In this work, we explore the spin-based sensing of random noises and show that the ambiguity problem is absent, while the dependence of $\mathcal{F}(\zeta)$ and $F(\zeta)$ on $\zeta_{true}$ makes adaptive measurements critical for improving the sensing precision.

Our subsequent discussions are based on the general adaptive measurement protocol in Fig. 1 which involve many important concepts and techniques, such as the QFI, the CFI, optimal unbiased estimators (such as the Bayesian estimator and the maximum likelihood estimator), the Cramér-Rao bound, and adaptive measurements. A systematic, self-contained introduction to these concepts (including a simple example) are given in Appendices A-E.

III. ADAPTIVE SENSING OF MAGNETIC NOISES

The main purpose of this section is to identify the distinguishing features of spin-based noise sensing compared with the estimation of deterministic Hamiltonian parameters. For this purpose, we consider a generic pure-dephasing model

$$\hat{H}(t) = [\omega + \tilde{\omega}(t)] \hat{S}_z, \quad (4)$$

describing the evolution of a spin-1/2 $\hat{S}$ under a constant magnetic field $B$ and a magnetic noise $\tilde{B}(t)$ along the $z$ axis, where $\omega \equiv \gamma B$, $\tilde{\omega}(t) \equiv \gamma \tilde{B}(t)$, and $\gamma$ is the gyromagnetic ratio. This model is relevant to many experiments involving a localized electron spin in solid state environments (such as the semiconductor quantum dots and nitrogen-vacancy centers in diamond), where the dominant magnetic noises come from the surrounding electron spin bath or nuclear spin bath. The former can be modelled by a Ornstein–Uhlenbeck noise [41, 73–75], while the latter can be modelled by a quasi-static noise [76, 77] (see Ref. [78] for a review). Next, we follow the standard steps outlined in Sec. II and further detailed in Appendices A-E to discuss the estimation of the noise $\tilde{\omega}(t)$ in comparison with the estimation of $\omega$ – a paradigmatic Hamiltonian parameter.

For step 1, we prepare the spin into a pure initial state

$$|\psi_{in}\rangle = \cos \frac{\Theta}{2} |\uparrow\rangle + \sin \frac{\Theta}{2} |\downarrow\rangle \quad (5)$$

parametrized by $\Theta$. Next, under the Hamiltonian in Eq. (4), the spin evolves for an interval $\tau$ into a final mixed state

$$\hat{\rho}(\tau) = \left( e^{i\omega \tau} e^{-\chi} \frac{\cos^2 \Theta}{2} \cos \frac{\Theta}{2} e^{i\omega \tau} e^{-\chi} \sin \frac{\Theta}{2} \cos \frac{\Theta}{2} e^{i\omega \tau} e^{-\chi} \sin \frac{\Theta}{2} \cos \frac{\Theta}{2} \right), \quad (6)$$

where $e^{-\chi} \equiv \langle e^{-i \int_{\tau}^{0} \tilde{\omega}(t) dt} \rangle$ is the average of the random phase over the noise distribution. For general noises, $\chi$ could be complex. Here we assume $\tilde{\omega}(t)$ is symmetric about zero, then $\chi$ is
real. Usually the fluctuation of the random phase grows with the evolution time $\tau$, so $e^{-\chi}$ decreases with $\tau$, corresponding to the decay of the average spin in the $xy$ plane or spin decoherence for short [78]. From Eq. (5), we see that all the information about $\omega$ is carried by the phase factor $e^{-i\omega\tau}$, while all the information about the noise is carried by the decoherence factor $e^{-\chi}$.

For $\omega$ and any parameter (denoted by $\xi$) that characterizes the noise $\tilde{\omega}(t)$, the QFI in the final state $\tilde{\rho}(\tau)$ can be computed by Eq. (A4) as

$$F_\omega = \tau^2 e^{-2\chi} \sin^2 \Theta,$$
$$F_\xi = (\partial_\xi \chi)^2 e^{-2\chi} \sin^2 \Theta.$$ (7a)

Here $F_\omega$ and $F_\xi$ show very different dependences on the decoherence factor $\chi$. This highlights the first distinguishing feature of noise sensing compared with the estimation of a deterministic magnetic field. For estimating $\omega$ ($\xi$), we should maximize $F_\omega$ ($F_\xi$) by tuning the controlling parameters $\Theta$ and $\tau$. The optimal value of $\Theta$ is $\Theta = \pi/2$. The optimal value of $\tau$ should be chosen to maximize $F_\omega$ ($F_\xi$). The optimal $\tau$ depends on the specific form of $\chi$ as a function of $\tau$, which in turn is determined by the details of the noise (to be discussed shortly).

For step 2, we consider a general projective measurement on the spin-1/2 along an axis with polar angle $\theta$ and azimuth $\varphi$. This measurement on $\tilde{\rho}(\tau)$ gives an outcome $\pm 1$ according to the probability distribution

$$P_{\pm 1} = \frac{1 \pm e^{-\chi} \sin \theta \cos (\omega \tau - \varphi)}{2}.$$ (8)

Here $P_{\pm 1}$ as a function of $\omega$ has a period $2\pi/\tau$, thus the measurement cannot distinguish $\omega$ and $\omega + 2n\pi/\tau$ ($n \in \mathbb{Z}$). This is the commonly encountered modulo-2$\pi$ ambiguity problem in Hamiltonian parameter estimation. By contrast, $P_{\pm 1}$ are usually not periodic in the noise parameters, so the modulo-2$\pi$ ambiguity is absent. This highlights the second distinguishing feature of noise sensing.

Given the measurement distribution, we can compute the CFI from Eq. (B1) and obtain

$$F_\omega = \tau^2 \frac{\sin^2 \theta \sin^2 (\omega \tau - \varphi)}{e^{2\chi} - \sin^2 \theta \cos^2 (\omega \tau - \varphi)},$$

$$F_\xi = (\partial_\xi \chi)^2 \frac{\sin^2 \theta \cos^2 (\omega \tau - \varphi)}{e^{2\chi} - \sin^2 \theta \cos^2 (\omega \tau - \varphi)}.$$ (9a)

We set $\theta = \pi/2$ and $\varphi = \omega \tau + \pi/2$ ($\varphi = \omega \tau$), so the CFI attains the QFI: $F_\omega = F_\omega$ ($F_\xi = F_\xi$). Namely, the optimal measurement for estimating $\omega$ ($\xi$) is along an axis in the $xy$ plane perpendicular (parallel) to the spin in the final state, as shown in Fig. 2. This highlights the third distinguishing feature of noise sensing.

For step 3, we adopt the maximum likelihood estimator (see Appendix C for details and Appendix D for an example) and leave the detailed numerical simulation to the next section.

Finally, we discuss how to optimize the evolution time $\tau$ to maximize the QFI $F_\omega = \tau^2 e^{-2\chi}$ and $F_\xi = (\partial_\xi \chi)^2/(e^{2\chi} - 1)$. For any classical noise (including static noises [80] and dynamical ones, Markovian noises and non-Markovian ones [78, 80]), once the statistics of the noise $\tilde{\omega}(t)$ is given, we can determine $\chi$ and hence $F_\omega$ and $F_\xi$ as functions of $\tau$ for any classical noise, at least in principle. Thus the method described here can be used to sensing an arbitrary classical noise, such as the abnormal static noises due to disorder averaging [80]. Moreover, although the discussions above are restricted to a single spin-1/2 (or equivalently a qubit), the method can also be used to infer the properties of noises on a general quantum system [80]. Compared with the spin-1/2 case, the difference is that the QFI should be calculated from Eq. (14) and the optimal measurement capable of converting all the QFI contained in the final density matrix of a general quantum system into the CFI is more complicated (see Appendix E).

Here for specificity we consider a widely used noise responsible for spin decoherence in electron spin baths [73–75]: the Ornstein–Uhlenbeck noise, which is a Gaussian noise characterized by the auto-correlation function

$$\langle \tilde{\omega}(t) \tilde{\omega}(t') \rangle = b^2 e^{-|t-t'|/\tau_c},$$

with $b$ ($\tau_c$) the amplitude (memory time) of the noise. The Wick’s theorem for Gaussian noises gives [78]

$$\chi = b^2 \tau_c^2 \left( \frac{\tau}{\tau_c} + e^{-\tau/\tau_c} - 1 \right) \approx \begin{cases} \frac{1}{2} b^2 \tau_c^2 & (\tau \ll \tau_c) \\ \left( b^2 \tau_c \right) \tau & (\tau \gg \tau_c). \end{cases}$$ (10)

For $\tau \ll \tau_c$, the spin decoherence is Gaussian: $e^{-\chi} = e^{-b^2 \tau_c^2/2}$. For $\tau \gg \tau_c$, the spin decoherence is exponential $e^{-\chi} \approx e^{-\tau/\tau_c}$ on a time scale

$$T_\varphi \equiv \frac{1}{b^2 \tau_c}.$$ (11)

Substituting Eq. (10) into Eq. (7) gives the QFI’s about $\omega, b,$
and \( \tau_c \), respectively:

\[
\mathcal{F}_\omega = \tau^2 e^{-2\chi},
\]

\[
\mathcal{F}_b = \frac{4\chi^2}{b^2 e^{2\chi} - 1},
\]

\[
\mathcal{F}_{\tau_c} = \frac{g^2(\tau/\tau_c)}{\tau_c^2} \chi^2 e^{2\chi} - 1,
\]

where

\[
g(x) \equiv \frac{x - 2 + e^{-x}(x + 2)}{x + e^{-x} - 1} \approx \begin{cases} \frac{1}{2} & (x \ll 1) \\ 1 & (x \gg 1) \end{cases}
\]

increases monotonically with \( x \) till saturation. With increasing evolution time \( \tau \), the decoherence factor \( \chi \) increases monotonically, so all the QFI’s first increases for small spin decoherence and then begin to decrease when the spin decoherence becomes significant. For \( \tau \ll \tau_c \), the QFI’s are given by

\[
\mathcal{F}_\omega = \tau^2 e^{-\beta\tau^2},
\]

\[
\mathcal{F}_b = \frac{b^2 \tau^4}{e^{b^2 \tau^2} - 1},
\]

\[
\mathcal{F}_{\tau_c} = \frac{1}{36} \frac{b^4 \tau^2}{e^{b^2 \tau^2} - 1} \tau^4.
\]

For \( \tau \gg \tau_c \), the QFI’s are given by

\[
\mathcal{F}_\omega = \tau^2 e^{-2\chi/\tau_c},
\]

\[
\mathcal{F}_b = 4b^2 \tau^2 e^{-2\chi/\tau_c} - 1,
\]

\[
\mathcal{F}_{\tau_c} = b^4 \tau^2 e^{2\chi/\tau_c} - 1.
\]

Next, we discuss the the optimization of the evolution time \( \tau \) to maximize the QFI.

For Markovian noises (i.e., \( bt_c \ll 1 \)), the spin coherence decays exponentially \( e^{-x} \approx e^{-\chi/\tau_c} \) on a time scale \( \tau_c \), as shown in Fig. 3(a). Thus \( \mathcal{F}_\omega \) is well approximated by Eq. (12a), shown as the black dotted line in Fig. 3(b). With increasing evolution time \( \tau \), \( \mathcal{F}_\omega \) first increases quadratically and then decays exponentially. At the optimal evolution time \( \tau_{\omega,\text{opt}} = T_\varphi \), it reaches the maximum \( \mathcal{F}_{\omega,\text{opt}} = T_\varphi^2 / e^2 \), as shown in Fig. 3(b). For noise sensing, \( \mathcal{F}_b \) and \( \mathcal{F}_{\tau_c} \) as functions of \( \tau \) differ from that of \( \mathcal{F}_\omega \) in that they exhibit three stages [Fig. 3(b)]. For \( \tau \ll \tau_c \), we have \( \mathcal{F}_b \simeq \tau^2 \) and \( \mathcal{F}_{\tau_c} \simeq \tau^4 \). For \( \tau_c \ll \tau \ll T_\varphi \), the evolution times of the QFI’s increase linearly with \( \tau \). For \( \tau \gg T_\varphi \), \( \mathcal{F}_b \) and \( \mathcal{F}_{\tau_c} \) decays exponentially with \( \tau \). At the optimal evolution time \( \tau_{\omega,\text{opt}} \approx \tau_{\varphi,\text{opt}} \approx 0.8T_\varphi \), \( \mathcal{F}_b \) and \( \mathcal{F}_{\tau_c} \) attain their maxima \( \mathcal{F}_{b,\text{opt}} \approx 0.65/b^2 \) and \( \mathcal{F}_{\tau_c,\text{opt}} \approx 0.162/\tau_c^2 \). For estimating \( \omega \), the optimal evolution time \( \tau_{\omega,\text{opt}} \) is independent of \( \omega \), thus adaptive measurements are not necessary. By contrast, for estimating the noise parameter \( b (\tau_c) \), the optimal evolution time \( \tau_{b,\text{opt}} (\tau_{\tau_c,\text{opt}}) \) depend on the parameter \( b (\tau_c) \) to be estimated, so adaptive measurements are crucial.

For non-Markovian noises (i.e., \( bt_c \gtrsim 1 \)), the Gaussian decay \( e^{-x} = e^{-\chi^2/\tau_c^2} \) for the spin coherence and the QFI’s [Eq. (12)] becomes appreciable even in the short-time regime \( \tau \ll \tau_c \), as shown in Fig. 3(c) and (d). In general, the peak location of the QFI as a function of \( \tau \) depends on both \( \tau_c \) and \( b \), thus adaptive measurements are crucial for estimating \( b \) and \( \tau_c \), as opposed to the estimation of \( \omega \).

The general analysis in this section have identified a series of distinguishing features for the sensing of noises compared to the estimation of deterministic magnetic fields, including the different dependences of the QFI’s on the spin decoherence, the absence of the modulo-2\( \pi \) phase ambiguity, the different optimal measurement schemes, and the crucial role of adaptive measurements. In the next section, we perform numerical simulations to demonstrate the feasibility of adaptive measurements to improve the precision of noise sensing.

### IV. ADAPTIVE MEASUREMENT OF SPIN DECOHERENCE TIME

Here we consider the decoherence of a localized spin caused by the surrounding electron spin bath in the solid state environment. The electron spin bath can be modelled by a Markovian noise \( \chi \) and leads to exponential spin decoherence \( e^{-\chi} \approx e^{-\chi/\tau_c} \) on a time scale \( \tau_c \) [see Eq. (11)]. The spin decoherence time \( \tau_c \) is a key characteristics of spin qubits in solid state environments. Next, we consider the adaptive estimation of the spin decoherence time \( \tau_c \). The rapid estimation of the spin decoherence time is not only important for the experimental characterization of the spin decoherence, but also allows us to design efficient coherence protection schemes to suppress the decoherence. Since \( \tau_c \) is a noise parameter, its adaptive estimation differs significantly from the estimation of a Hamiltonian parameter.

According to Sec. III the optimal initial state of the spin is Eq. (5) with \( \Theta = \pi/2 \). In the following we consider two different protocols to measure the spin decoherence time \( \tau_c \):

![Image of graphs showing spin coherence and quantum Fisher information as functions of evolution time \( \tau \).](image-url)
FIG. 4. Classical Fisher information as a function of the evolution time for (a) spin echo protocol and (b) free evolution protocol with \(\omega T_\varphi = 400\pi/3 \gg 1\). The symbols denote \((\delta T_\varphi)^{-2}/N\), where \(\delta T_\varphi\) is the sensing precision from numerically simulating \(N = 10^3\) repeated measurements with different evolution time: red for \(\tau = 0.1T_\varphi\), green for \(\tau = 0.8T_\varphi\), blue for \(\tau = 3T_\varphi\). The inset of (b) zooms in on dashed square region.

(i) Spin echo; (ii) Free evolution.

For spin echo, we first let the spin evolve under the Hamiltonian in Eq. (14) for an interval \(\tau/2\), next apply an instantaneous \(\pi\)-pulse to induce the flip between \(|\uparrow\rangle\) and \(|\downarrow\rangle\), and finally let the spin evolve under the Hamiltonian in Eq. (14) for another interval \(\tau/2\) into the final state \(\hat{\rho}(\tau) = (1+e^{-i\tau/2}\hat{\sigma}_z)/2\). The QFI about \(T_\varphi\) in this final state is

\[
\mathcal{F}(\tau) = \frac{\tau^2}{T_\varphi^2}\frac{1}{e^{2\pi\tau T_\varphi} - 1},
\]

(14)

where we have made its dependence on the evolution time \(\tau\) explicit. This spin echo technique [81] can eliminate quasi-static noises (such as those from the surrounding nuclear spins of the host lattice) and single out the decoherence caused by the Markovian noise under consideration. Interestingly, it also eliminates the Larmor frequency \(\omega\), so that the final density matrix \(\hat{\rho}(\tau)\) is independent of \(\omega\). Finally, we perform a projective measurement on the spin along the \(x\) axis. According to Sec. III, this measurement is optimal. Indeed, it gives an outcome \(\pm 1\) according to the probability distribution

\[
P(\pm 1|T_\varphi) = \frac{1 \pm e^{-i\tau T_\varphi}}{2},
\]

(15)

and the CFI in each outcome attains the QFI:

\[
F(\tau) = \frac{\tau^2}{T_\varphi^2}\frac{1}{e^{2\pi\tau T_\varphi} - 1},
\]

(16)

as shown in Fig. 4(a).

For free evolution, we simply let the spin evolve under the Hamiltonian in Eq. (4) for an interval \(\tau\) into the final state \(\hat{\rho}(\tau) = 1/2 + (e^{-i\tau T_\varphi}/2)[\hat{\sigma}_z \cos(\omega \tau) + \hat{\sigma}_x \sin(\omega \tau)]\), i.e., Eq. (6) with \(\Theta = \pi/2\) and \(\chi = \tau T_\varphi\). This final state differs from that of the spin echo protocol in that it still depends sensitively on the Larmor frequency \(\omega\). Consequently, in order to measure \(T_\varphi\) from this free-evolution final state, precise knowledge about \(\omega\) is usually necessary (to be discussed shortly), although the QFI about \(T_\varphi\) contained in this final state is still given by Eq. (14), i.e., the same as the spin echo protocol. Finally, we should perform an optimal measurement to convert all the QFI into the CFI. According to Sec. III, the optimal measurement is a projective one along the azimuth \(\omega \tau\) in the \(xy\) plane (dashed blue line in Fig. 2), whose CFI is equal to the QFI in Eq. (14). However, in our adaptive measurement scheme (to be discussed shortly), the parameter \(\tau\) and hence the measurement axis will vary in different measurement cycles. The frequent change of the measurement axis may complicates its experimental realization. To avoid this problem, we fix the measurement axis to be along the \(x\) axis (i.e., we always measure \(\hat{\sigma}_x\)), then the measurement distribution is

\[
P(\pm 1|T_\varphi) = \frac{1 \pm e^{-i\tau T_\varphi} \cos(\omega \tau)}{2},
\]

(17)

and the CFI in each outcome,

\[
F(\tau) = \frac{\tau^2}{T_\varphi^4} \frac{\cos^2(\omega \tau)}{e^{2\pi\tau T_\varphi} - \cos^2(\omega \tau)},
\]

(18)

shows rapidly oscillation as a function of \(\tau\), with its envelope coinciding with the QFI [see Fig. 4(b)]. In other words, the CFI still attains the QFI when \(\tau\) is an integer multiple of \(\pi/\omega\), but does not attains the QFI for general \(\tau\). Fortunately, we can still tune \(\tau\) to maximize the QFI and the CFI simultaneously.

Now we optimize the evolution time \(\tau\). For spin echo, the optimal \(\tau\) is [see Fig. 4(a)]

\[
\tau_{\text{opt}} = 0.8T_\varphi,
\]

(19)

For free evolution, under the realistic assumption \(\omega \gg 1/T_\varphi\), the optimal \(\tau\) is an integer multiple of \(\pi/\omega\) closest to \(0.8T_\varphi\) [see Fig. 4(b)]:

\[
\tau_{\text{opt}} = \frac{n\pi}{\omega} (n \in \mathbb{Z}) \quad \text{and} \quad \tau_{\text{opt}} \approx 0.8T_\varphi,
\]

(20)

For both the spin echo protocol and the free evolution protocol, choosing \(\tau = \tau_{\text{opt}}\) gives the same maximal CFI and QFI,

\[
F_{\text{opt}} = \mathcal{F}_{\text{opt}} = \frac{0.16}{T_\varphi},
\]

(21)

and hence the same optimal sensing precision

\[
(\Delta T_\varphi)_{\text{opt}} = \frac{2.5T_\varphi}{\sqrt{N}}
\]

(22)

for \(N\) repeated measurements. The difference is that in the spin echo (free evolution) protocol, \(\tau_{\text{opt}}\) is independent of (dependent on) the Larmor frequency \(\omega\). Specifically, in the free evolution protocol, the rapid oscillation of the CFI as a function of \(\tau\) with a period \(\pi/\omega\) [see Fig. 4(b)] requires precise knowledge about \(\omega\) and high control precision of \(\tau\) on the order of \(1/\omega\) to correctly locate the maximum [i.e., Eq. (20)] of the CFI. By contrast, in the spin echo protocol, the CFI as a function of \(\tau\) is independent of \(\omega\) [see Fig. 4(a)], so it requires no knowledge about \(\omega\) and relatively low control precision of \(\tau\) on the order of \(1/T_\varphi \gg 1/\omega\) to correctly locate the maximum of the CFI.

Unfortunately, for both protocols, \(\tau_{\text{opt}}\) depends on the unknown parameter \(T_\varphi\). Due to this dependence, adaptive
schemes that update \( \tau_{\text{opt}} \) after each measurement cycle can outperform significantly non-adaptive ones. For each protocol, we consider three different measurement schemes involving different treatments of the evolution time \( \tau \): repeated measurements, adaptive measurements, and the least-square fitting that is commonly used in experiments.

### A. Repeated measurement scheme

The evolution time \( \tau \) is fixed during the entire estimation process. After repeating the initialization-evolution-measurement cycle \( N \) times, we get \( N \) outcomes \( \mathbf{u} = (u_1, u_2, \cdots, u_N) \). Using these outcomes, we refine our knowledge about \( T_\psi \) to the posterior distribution

\[
P_u(T_\psi) \sim \left[ P(1|T_\psi) \right]^{N_u} \left[ P(-1|T_\psi) \right]^{N_{-1}},
\]

where \( N_u \) (\( N_{-1} \)) is the number of outcome +1 (−1) and \( P(±1|T_\psi) \) is given by Eq. (15) for spin echo and Eq. (17) for free evolution. Finally, we construct the maximum likelihood estimator \( T_M \equiv \arg \max P_u(T_\psi) \) and quantify its precision by [cf. Eq. (C3)]

\[
\delta T_\psi \equiv \sqrt{\int (T_M - T_\psi)^2 P_u(T_\psi) dT_\psi}.
\]

To analyze the performance of this scheme, we notice that for large \( N \), the maximum likelihood estimator is known to be unbiased and can saturate the Cramér-Rao bound Eq. (3), so the sensing precision can be approximated by

\[
(\delta T_\psi)_{\text{CRB}} \equiv \frac{1}{\sqrt{NF(\tau)}}.
\]

Here the CFI \( F(\tau) \) is given by Eq. (16) for spin echo and Eq. (18) for free evolution (see Fig. 4). The evolution time \( \tau \) directly determines the sensing precision, e.g., setting \( \tau = \tau_{\text{opt}} \) would lead to the optimal sensing precision in Eq. (22). However, \( \tau_{\text{opt}} \) is unknown because it depends on the unknown parameter \( T_\psi \) to be estimated. This makes adaptive measurements crucial for achieving the optimal sensing precision.

### B. Adaptive measurement schemes

The key idea is to use the outcomes of previous measurement to refine our knowledge about \( T_\psi \) and then use this knowledge to optimize \( \tau \). We consider two different adaptive schemes: the CFI-based scheme [82–84] and the locally optimal adaptive scheme [63, 64, 85, 86], as introduced in Appendix E. The former updates the maximum likelihood estimator \( T_M \) after every initialization-evolution-measurement cycle and then set the evolution time to

\[
\tau = 0.8T_M
\]

for the spin echo protocol or

\[
\tau = \frac{n}{\omega} \quad \text{and} \quad \tau \approx 0.8T_M \quad (n \in \mathbb{Z})
\]

for the free evolution protocol. The latter optimizes \( \tau \) to minimize the expected uncertainty of the estimator at the end of the next cycle (see Appendix E). Suppose at the end of the \((n - 1)\)th cycle, our knowledge about \( T_\psi \) is quantified by the distribution \( P(T_\psi) \) and the maximum likelihood estimator \( T_M \equiv \arg \max P(T_\psi) \) constructed from the outcomes of all the previous cycles. In the \( n \)th cycle with the evolution time \( \tau \), the measurement distribution \( P(±1|T_\psi) \) [Eq. (15) or Eq. (17)] depends on \( T_\psi \) and \( \tau \). If the outcome is \( u \), then our knowledge would be refined to \( P_u(T_\psi) \sim P(T_\psi)P(u|T_\psi) \), which in turn gives the maximum likelihood estimator \( T_M(u, \tau) \equiv \arg \max P_u(T_\psi) \) and its uncertainty [cf. Eq. (C3)]

\[
\delta T_\psi(u, \tau) \equiv \sqrt{\int (T_M - T_\psi)^2 P_u(T_\psi)dT_\psi}.
\]

Since the probability for outcome \( u \) is estimated as \( P(u|T_M) \), we would choose \( \tau \) in the \( n \)th cycle to minimize the expected uncertainty [cf. Eq. (25)]

\[
\overline{\delta T_\psi(\tau)} = \sum_{u=±1} P(u|T_M)\overline{\delta T(u, \tau)}.
\]

For \( n = 1 \), i.e., the first cycle, there is no prior information, i.e., \( P(T_\psi) \) is a constant, so \( T_M \) is chosen randomly.

To analyze the performance, we notice that after a large number of adaptive steps, the estimator \( T_M \) would approach the true decoherence time \( T_\psi \). Consequently, according to Appendix E, the evolution time \( \tau \) and hence the sensing precision for these two adaptive schemes would coincide with each other. In addition, the evolution time \( \tau \) in Eqs. (25) and (26) would approach \( \tau_{\text{opt}} \), so the corresponding sensing precision would approach the optimal precision \( (\delta T_\psi)_{\text{opt}} \) in Eq. (22).

### C. Least-square fitting scheme

For a given range \([0, \tau_{\text{max}}]\) of the evolution time, we uniformly discretize it into \( M \gg 1 \) grids \( \tau_k = k\Delta \tau \), where \( \Delta \tau = \tau_{\text{max}}/M \). For each \( \tau_k \), we repeat the \( \hat{\sigma}_z \) measurement \( \nu \) times and calculate their average. Then we fit this average as a function of \( \tau \) to the theoretical curve \( \langle \hat{\sigma}_z \rangle = e^{-\tau/T_\psi} \) (for the spin echo protocol) or \( \langle \hat{\sigma}_x \rangle = \cos(\omega \tau)e^{-\tau/T_\psi} \) (for the free evolution protocol) to obtain an estimator to \( T_\psi \). Finally, we repeat the procedures above for \( q \gg 1 \) times to obtain many estimators \( \tilde{T}_{\psi,\text{est}}^{(1)}, \tilde{T}_{\psi,\text{est}}^{(2)}, \cdots, \tilde{T}_{\psi,\text{est}}^{(q)} \) and determine the uncertainty \( \delta T_\psi \) of a single estimator as the square root of the statistical variance these estimators, i.e.,

\[
\delta T_\psi \equiv \frac{1}{q} \sum_{i=1}^q (\tilde{T}_{\psi,\text{est}}^{(i)} - \bar{T}_{\psi,\text{est}})^2,
\]

where \( \bar{T}_{\psi,\text{est}} = (1/q) \sum_{i=1}^q T_{\psi,\text{est}}^{(i)} \).

According to the Cramér-Rao bound in Eq. (3), the sensing precision of this scheme can be roughly estimated as

\[
(\delta T_\psi)_{\text{CRB}} \equiv \frac{1}{\sqrt{NF_{\text{ave}}}}.
\]
adaptive sensing (green line) and locally optimal adaptive sensing (blue line). The black (gray) dotted line indicates \((\delta T_\psi)_\text{opt}\) in Eq. (22) ([\(\delta T_\psi\)]_\text{CRB} in Eq. (27)). (b) Successive refinement of the evolution time \(\tau\) in CFI-based (green line) and locally optimal (blue line) adaptive sensing.

where \(N \equiv \nu M\) is the total number of measurements,

\[
F_{\text{ave}} = \frac{1}{\tau_{\text{max}}} \int_0^{\tau_{\text{max}}} F(\tau) d\tau
\]

is the average of the CFI over the range \([0, \tau_{\text{max}}]\). For spin echo, \(F(\tau)\) in Eq. (16) decays exponentially for large \(\tau\) [see Fig. 4(a)]. For \(\tau_{\text{max}} \gg T_\psi\), we have \(F_{\text{ave}} \approx 0.3/(T_\psi \tau_{\text{max}})\), thus

\[
(\delta T_\psi)_{\text{CRB}} \approx 1.8 \sqrt{\frac{T_\psi \tau_{\text{max}}}{N}} \approx 0.7 \sqrt{T_{\text{max}} T_\psi} (\delta T_\psi)_{\text{opt}}
\]

(27)
degrades monotonically with increasing \(\tau_{\text{max}}\). For free evolution, \(F(\tau)\) in Eq. (15) shows rapid oscillations as a function of \(\tau\) with an envelope coinciding with the CFI for spin echo [see Fig. 4(b)]. For \(\tau_{\text{max}} \gg T_\psi\), we have \(F_{\text{ave}} \approx 0.14/(T_\psi \tau_{\text{max}})\), which is about half that of the spin echo, thus \((\delta T_\psi)_{\text{CRB}} \approx \sqrt{T_{\text{max}} T_\psi} (\delta T_\psi)_{\text{opt}}\) also degrades with increasing \(\tau_{\text{max}}\).

D. Numerical simulations

In the above, we have presented two protocols to measure the spin decoherence time \(T_\psi\): the spin echo protocol and the free evolution protocol. For each protocol, we consider three kinds of schemes, which involve different treatments of the evolution time \(\tau\): (i) The repeated measurement scheme uses a fixed \(\tau\); (ii) The two adaptive measurement schemes update \(\tau\) in every measurement cycle; (iii) The least-square fitting scheme scan \(\tau\) over a fixed range. The spin echo protocol singles out the \(T_\psi\) process from all the other unwanted evolution, so the measurement distribution \(P(\pm 1|T_\psi)\) is independent of the Larmor frequency \(\omega\). Consequently, all the schemes applied to the spin echo protocol require no knowledge about \(\omega\) and a relatively low control precision (on the order of \(1/T_\psi\)) over \(\tau\). By contrast, the free evolution protocol leaves the Larmor precession intact, so the measurement distribution \(P(\pm 1|T_\psi)\) [see Eq. (17)] depends sensitively on \(\omega\). Consequently, all the schemes applied to this protocol require precise knowledge about \(\omega\) and much higher (on the order \(1/\omega\)) control precision over \(\tau\). Specifically: (i) In the repeated measurement scheme, since the posterior distribution \(P_\nu(T_\psi)\) depends on \(\omega\), we cannot find the maximum likelihood estimator \(T_M \equiv \arg \max P_\nu(T_\psi)\) if \(\omega\) is unknown; (ii) In the CFI-based adaptive scheme, we cannot set \(\tau\) to Eq. (26) if \(\omega\) is unknown; in the locally optimal adaptive scheme, the expected uncertainty \(\delta T_\psi(\tau)\) depend on both \(\tau\) and \(\omega\), so we cannot find the minimum of \(\delta T_\psi(\tau)\) as a function of \(\tau\) if \(\omega\) is unknown. (iii) In the least-square fitting scheme, it would be difficult to choose the grid spacing \(\Delta \tau\) and to fit the measurement data to \((\delta \hat{\tau}) = \cos(\omega \tau) e^{-\tau/T_\psi}\) to extract \(T_\psi\) when \(\omega\) is unknown. Therefore, the spin echo protocol is advantageous over the free evolution protocol if our knowledge about \(\omega\) is limited or the available control precision over \(\tau\) is low.

In all our numerical simulations, we take the true value \(T_\psi\) as the unit of time, i.e., \(T_\psi = 1\), and take the true value of \(\omega\) to be \(\omega = 400\pi/3 \gg 1/T_\psi\).

To begin with, we check the sensing precision of the repeated measurement scheme applied to the spin echo protocol and the free evolution protocol. We consider three sets of evolution time: \(\tau = 0.1, 0.8,\) and \(3\). For each case, our numerical simulations show that with increasing number \(N\) of repeated measurements, the uncertainty \(\delta T_\psi\) calculated from Eq. (23) gradually approaches the large-\(N\) limit \((\delta T_\psi)_{\text{CRB}}\) [Eq. (24)] for both protocols. For example, the uncertainty \(\delta T_\psi\) for \(N = 10^4\) repeated measurements agree well with \((\delta T_\psi)_{\text{CRB}}\), i.e., \((\delta T_\psi)^2/N\) agree well with the CFI \(F(\tau)\), as shown in Fig. 4(a) and 4(b).

Next, we consider the spin echo protocol and compare the adaptive scheme with the commonly used least-square fitting scheme with \(\tau_{\text{max}} = 10\) and \(\nu = q = 100\). As shown in Fig. 5(a), the precision of the least-square fitting is well approximated by \((\delta T_\psi)_{\text{CFI}}\) in Eq. (27), which is significantly worse than the optimal sensing precision \((\delta T_\psi)_{\text{opt}}\) in Eq. (24). By contrast, the precision of both the CFI-based adaptive scheme and the locally optimal adaptive scheme approaches the optimal precision \((\delta T_\psi)_{\text{opt}}\) after \(\approx 100\) measurements. Physically, this is because both adaptive schemes successively adjust the evolution time [e.g., Eqs. (25) and (26) for the CFI-based adaptive scheme and the free evolution protocol] based on the newest knowledge about the unknown parameter \(T_\psi\) after every measurement. As shown in Fig. 5(b), after \(\sim 100\) measurements, the evolution time in both adaptive schemes already approaches the optimal evolution time \(T_{\psi\text{opt}}\) [Eq. (19)].

Then, we turn to the free evolution protocol, which requires precise knowledge about \(\omega\) and precise control over \(\tau\) on the order \(1/\omega\). Since the average spin \((\hat{\tau})_r = \cos(\omega \tau) e^{-\tau/T_\psi}\) exhibits rapid oscillations as functions of the evolution time \(\tau\), the best way to do least-square fitting is to let the grid spacing \(\Delta \tau\) be an integer multiple of \(\pi/\omega\), so that \(\tau_r = k \Delta \tau\) samples the envelope of the \((\hat{\tau})_r\) curve only. This also amounts to sampling the envelope of the rapidly oscillating CFI in Eq. (18), so that \(F(\tau_r) = F(\tau_k)\) attains the corresponding CFI.

Therefore, the least-square fitting scheme applied to the free evolution protocol would give the same precision as it does for the spin echo protocol, so we do not simulate this case any more. For the two adaptive schemes applied to the free evolution protocol, as shown in Fig. 6(a), both the CFI-based one and the locally optimal one approach the optimal sensing.
FIG. 6. Numerical simulation for the estimation of the spin decoherence time $T_\phi$ by free evolution. (a) Estimation precision $\delta T_\phi$ by CFI-based adaptive sensing (green line) and locally optimal adaptive sensing (blue line). The black dotted line indicates the optimal sensing precision $\delta T_\phi^{\text{opt}}$ in Eq. (23). (b) Successive refinement of the evolution time $t$ in CFI-based (green lines) and locally optimal (blue lines) adaptive sensing.

precision $\delta T_\phi^{\text{opt}}$ after $\sim 100$ measurements, similar to the case of the spin echo protocol. As shown in Fig. 6b, the adaptive schemes successively refine the evolution time [e.g., Eq. (26)] based on the newest knowledge about $T_\phi$ after every measurement. After $\sim 100$ measurements, the evolution time in both adaptive schemes approach the optimal evolution time $\tau_{\text{opt}}$ [Eq. (20)].

V. CONCLUSION

Using localized spins as ultrasensitive quantum sensors is attracting widespread interest. Recently, adaptive measurements were used to improve the dynamic range for the spin-based estimation of deterministic Hamiltonian parameters such as the external magnetic field. Here we explore a very different direction – the use of adaptive measurements in spin-based sensing of random noises. We have performed general analysis that identifies a series of important differences between noise sensing and the estimation of deterministic magnetic fields, such as the different dependences on the spin decoherence, the different optimal measurement schemes, the absence of the modulo-$2\pi$ phase ambiguity, and the crucial role of adaptive measurement. We have also performed numerical simulations that clearly demonstrate significant speed up of the characterization of the spin decoherence time via adaptive measurements compared with the commonly used least-square fitting method. This work paves the way towards adaptive noise sensing.

ACKNOWLEDGMENTS

This work was supported by the MOST of China (Grants No. 2014CB848700), the National Key R&D Program of China (Grants No. 2017YFA0303400), the NSFC (Grants No. 11774021), and the NSFC program for “Scientific Research Center” (Grant No. U1530401). We acknowledge the computational support from the Beijing Computational Science Research Center (CSRC).

Appendix A: State preparation and encoding: quantum Fisher information

The amount of information about $\zeta$ contained in a general $\zeta$-dependent quantum state $\hat{\rho}_\zeta$ is quantified by its QFI [70]

$$F \equiv \text{Tr} \hat{\rho}_\zeta \hat{L}_\zeta^2,$$

(A1)

where $\hat{L}_\zeta$ is the so-called symmetric logarithmic derivative operator: it is an Hermitian operator defined through [69]

$$\partial_\zeta \hat{\rho}_\zeta = \frac{1}{2} (\hat{L}_\zeta \hat{\rho}_\zeta + \hat{\rho}_\zeta \hat{L}_\zeta).$$

The QFI defined in Eq. (A1) remains invariant under any $\zeta$-independent unitary transformations, i.e., such transformations conserves the quantum information. For a pure state $\hat{\rho}_\zeta = |\Phi\rangle\langle\Phi|$, we have $\hat{L}_\zeta = 2\partial_\zeta \hat{\rho}_\zeta$ and hence

$$F = 4 \langle (\partial_\zeta |\Phi\rangle |\partial_\zeta |\Phi\rangle \rangle^2 - 4 G_{\text{rms}}^2,$$

(A2)

where the last step applies to unitary evolution $|\Phi\rangle = e^{-iG_{\text{rms}}}|\Phi_{\text{ini}}\rangle$ and $G_{\text{rms}} = \langle (|\Phi_{\text{ini}}\rangle \langle G_{\text{rms}}^2 |\Phi_{\text{ini}}\rangle - (|\Phi_{\text{ini}}\rangle \langle |\Phi_{\text{ini}}\rangle)^2 \rangle)^{1/2}$ is the root-mean-square fluctuation of $G$ in the initial state. For a general mixed state with the spectral decomposition $\hat{\rho}_\zeta = \sum_n p_n |\Phi_n\rangle \langle\Phi_n|$, its QFI is [87–89]

$$F = \sum_n \frac{(\partial_\zeta p_n)^2}{p_n} + \sum_n p_n F_n - \sum_m \sum_n \frac{8 p_m p_n}{p_m + p_n} \langle (|\Phi_m\rangle \langle \partial_\zeta |\Phi_n\rangle)^2 \rangle,$$

(A3)

where $p_n$ are nonzero eigenvalues of $\hat{\rho}_\zeta$, $|\Phi_n\rangle$ are the corresponding ortho-normalized eigenstates, and $F_n$ is the QFI of the pure state $|\Phi_n\rangle$ [see Eq. (A2)]. This expression shows that the QFI of a non-full-rank state is completely determined by its support, i.e., the subset of $|\Phi_n\rangle \langle\Phi_n|$ with nonzero eigenvalues. For a two-level system, its density matrix can always be expressed in terms of the Pauli matrices $\hat{\sigma}$ as $\hat{\rho}_\zeta = (1/2)(1 + \hat{\sigma} \cdot n)$, where $n = \text{Tr} \hat{\rho}_\zeta \hat{\sigma}$ is the Bloch vector. The QFI for such a state is [90–92]

$$F = |\partial_\zeta n|^2 + \frac{(\partial_\zeta |n|^2)^2}{4(1 - |n|^2)},$$

(A4)

where the second term is absent when $|n| = 1$, i.e., when $\hat{\rho}_\zeta$ is a pure state. When $\hat{\rho}_\zeta = \hat{\rho}_\zeta^{(1)} \otimes \cdots \otimes \hat{\rho}_\zeta^{(N)}$ is the direct product state of $N$ quantum systems, its QFI is additive: $F = \sum_{n=1}^N F^{(n)}$, where $F^{(n)}$ is the QFI of $\hat{\rho}_\zeta^{(n)}$.

Physically, the QFI measures the rate of variation of $\hat{\rho}_\zeta$ with the parameter $\zeta$, e.g., if we regard $\hat{\rho}_\zeta$ and $\hat{L}_\zeta$ as classical variables, then $L_{\zeta} = \partial_\zeta \ln \hat{\rho}_\zeta$ and Eq. (A1) becomes the average of $(\partial_\zeta \ln \hat{\rho}_\zeta)^2$ over the state $\hat{\rho}_\zeta$. Moreover, the Bures distance between two quantum states $\hat{\rho}_1$ and $\hat{\rho}_2$ is defined as [93]

$$D(\hat{\rho}_1, \hat{\rho}_2) = \sqrt{2(1 - \text{Tr} \sqrt{\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_1 \hat{\rho}_2})^{1/2}},$$

where the second term $\text{Tr} \sqrt{\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_1 \hat{\rho}_2}$ on the right-hand side is the so-called Uhlmann fidelity [94]. For neighboring states $\hat{\rho}_\zeta$ and $\hat{\rho}_\zeta + d\zeta$, the Bures distance reduces to

$$D(\hat{\rho}_\zeta, \hat{\rho}_\zeta + d\zeta) = \frac{1}{2} \sqrt{d\zeta},$$
so the QFI measures the distinguishability between two neighboring states parametrized by \( \zeta \).

The importance of the QFI for parameter estimation is manifested in the inequalities Eqs. \((1)\) and \((3)\). Namely, given \( \hat{\rho}_\zeta \) and hence \( \mathcal{F}(\zeta) \), the precision of any unbiased estimator from \( N \) repetitions of any measurement is limited by the inequality

\[
\delta \zeta \geq \frac{1}{\sqrt{N \mathcal{F}(\zeta)}} \tag{A5}
\]

known as the quantum Cramér-Rao bound \([63, 70]\). Saturating this bound requires saturating Eqs. \((1)\) and \((3)\) simultaneously, i.e., using optimal measurements to convert all the QFI into the CFI and using optimal unbiased estimators to convert all the CFI into the precision of the estimator.

Appendix B: Measurement: classical Fisher information

A general measurement with discrete outcomes \( u \) is described by the positive-operator valued measure (POVM) elements \( \{M_u\} \) satisfying the completeness relation \( \sum_u M_u^\dagger M_u = 1 \). Given a quantum state \( \hat{\rho}_\zeta \), it yields an outcome \( u \) according to the probability distribution \( P(u|\zeta) \equiv \text{Tr} M_u^\dagger \hat{\rho}_\zeta M_u \) that depends on \( \zeta \). The amount of information about \( \zeta \) contained in each outcome is quantified by the CFI \([95]\):

\[
F(\zeta) \equiv \sum_u P(u|\zeta) \left( \frac{\partial \ln P(u|\zeta)}{\partial \zeta} \right)^2. \tag{B1}
\]

For continuous outcomes, we need only replace \( \sum_u \) by \( \int du \) everywhere. Physically, the CFI quantifies the dependence of the measurement distribution \( P(u|\zeta) \) on the parameter \( \zeta \). Actually, the Wootters’ distance \([96]\) between two probability distributions \( P^{(1)}(u) \) and \( P^{(2)}(u) \) is

\[
D(P^{(1)}, P^{(2)}) \equiv \cos^{-1} \left( \sum_u \sqrt{P^{(1)}(u)P^{(2)}(u)} \right).
\]

For neighboring distributions \( P^{(1)}(u) = P(u|\zeta) \) and \( P^{(2)}(u) = P(u|\zeta + d\zeta) \), the Wootters’ distance reduces to

\[
D(P(u|\zeta), P(u|\zeta + d\zeta)) = \frac{1}{2} \sqrt{F(\zeta)} d\zeta,
\]

so the CFI measures the distinguishability between neighboring measurement distributions parametrized by \( \zeta \).

Since the probability distribution function is the classical counterpart of the quantum mechanical density matrix, the CFI (Wootters’ distance) is the classical counterpart of the QFI (Bures distance). The inequality Eq. \((1)\) expresses the simple fact that no new information about \( \zeta \) can be generated in the measurement process: optimal (non-optimal) measurements convert all (part) of the QFI into the CFI. Given \( \hat{\rho}_\zeta \), the optimal measurement is not unique. The projective measurement on the symmetric logarithmic derivative operator \( L_{\text{sym}} \) has been identified \([70]\) as an optimal measurement, but \( \hat{L}_{\text{sym}} \) is not known. To circumvent this problem, the simplest way is to find other optimal measurements that do not depend on \( \zeta_{\text{true}} \). Another solution \([71]\) is to approximate \( L_{\text{sym}} \) by \( L_{\text{opt}} \), where \( \zeta_{\text{opt}} \) is our best guess to \( \zeta_{\text{true}} \), i.e., the optimal unbiased estimator, as we discuss below.

Appendix C: Data processing: optimal unbiased estimators

Given the measurement distribution \( P(u|\zeta) \) and hence the CFI \( F(\zeta) \) of each outcome, the precision \( \delta \zeta \) of any unbiased estimator \( \zeta_{\text{opt}}(u) \) constructed from the outcomes \( u \equiv (u_1, \cdots, u_N) \) of \( N \) repeated measurements is limited by the Cramér-Rao bound Eq. \((3)\), which expresses the simple fact that no new information about \( \zeta \) can be generated in the data processing: optimal (non-optimal) unbiased estimators convert all (part) of the CFI into the useful information \( (\delta \zeta)^{-2} \) quantified by the precision \( \delta \zeta \). Finding optimal unbiased estimators is an important step in parameter estimation. In the limit of large \( N \), two kinds of estimators are known to be unbiased and optimal: the maximum likelihood estimator and the Bayesian estimator \([95]\), as we introduce now.

Before any measurements, our prior knowledge about the unknown parameter \( \zeta \) is quantified by the posterior distribution \( P(\zeta) \), e.g., a \( \delta \)-like distribution corresponds to knowing \( \zeta \) exactly, a flat distribution corresponds to completely no knowledge about \( \zeta \), while a Gaussian distribution \( P(\zeta) \propto e^{-\langle \zeta \rangle^2/(2\sigma_0^2)} \) corresponds to knowing \( \zeta \) to be \( \zeta_0 \) with a typical uncertainty \( \sigma_0 \).

Upon getting the first outcome \( u_1 \), our knowledge about \( \zeta \) is immediately refined from \( P(\zeta) \) to

\[
P(\zeta|u_1) = \frac{P(\zeta)P(u_1|\zeta)}{N(u_1)}
\]

according to the Bayesian rule \([77]\), where \( N(u_1) \equiv \int d\zeta P(\zeta)P(u_1|\zeta) \) is a normalization factor ensuring \( P(\zeta|u_1) \) is normalized to unity: \( \int P(\zeta|u_1) d\zeta = 1 \). Here \( P(\zeta|u_1) \) is the posterior probability distribution of \( \zeta \) conditioned on the outcome of the measurement being \( u_1 \): its parametric dependence on \( u_1 \) means that different measurement outcomes leads to different refinement of knowledge about \( \zeta \).

Upon getting the second outcome \( u_2 \), our knowledge is immediately refined from \( P(\zeta|u_1) \) to

\[
P(\zeta|u_1, u_2) = \frac{P(\zeta)P(u_1|\zeta)P(u_2|\zeta)}{N(u_1, u_2)}
\]

where \( N(u_1, u_2) = \int P(\zeta)P(u_1|\zeta)P(u_2|\zeta) d\zeta \) is a normalization factor for the posterior distribution \( P(\zeta|u_1, u_2) \). If we omit the trivial normalization factors, then the measurement-induced knowledge refinement becomes

\[
P(\zeta) \overset{u_1}{\longrightarrow} P(\zeta)P(u_1|\zeta) \overset{u_2}{\longrightarrow} P(\zeta)P(u_1|\zeta)P(u_2|\zeta) \overset{\cdots}{\longrightarrow}.
\]

Upon getting \( N \) outcomes \( u \equiv (u_1, \cdots, u_N) \), our knowledge about \( \zeta \) is quantified by the posterior distribution

\[
P(\zeta) \sim P(\zeta)P(u|\zeta)
\]
up to a trivial normalization factor, where \( P(\mathbf{u}|\zeta) = P(u_1|\zeta) \cdots P(u_N|\zeta) \) is the probability for getting the outcome \( \mathbf{u} \). The posterior distribution \( P_u(\zeta) \) completely describe our state of knowledge about \( \zeta \). Nevertheless, sometimes a single number, i.e., an unbiased estimator, is required as the best guess to \( \zeta_{\text{true}} \). There are two well-known estimators: the maximum likelihood estimator \[95\]

\[ \zeta_{\text{M}}(\mathbf{u}) \equiv \arg \max_{\zeta} P_u(\zeta) \] (C1)

is the peak position of \( P_u(\zeta) \) as a function of \( \zeta \), while the Bayesian estimator \[98\]

\[ \zeta_{\text{B}}(\mathbf{u}) \equiv \int \xi P_u(\zeta)d\zeta \] (C2)

is the average of \( \zeta \). For large \( N \), both estimators are unbiased and optimal: \( \langle \zeta \rangle = \zeta \) and \( \delta \zeta_{\text{avg}} = 1/\sqrt{N F(\zeta)} \), where \( \alpha = \text{M or B} \), and \( (\cdot) \) denotes the average over a large number of estimators obtained by repeating the \( N \)-outcome estimator scheme many times and \( \delta \zeta_{\text{avg}} \) is defined as Eq. (2) or

\[ \delta \zeta_{\text{avg}} = \sqrt{\langle [\zeta - \zeta_{\text{avg}}(\mathbf{u})]^2 \rangle P_u(\zeta) d\zeta} \] (C3)

For a simple understanding, we consider \( N \to \infty \), so the number of occurrence of a specific outcome \( u \) approaches \( N P(u|\zeta_{\text{true}}) \). Then, up to a trivial normalization factor, the posterior distribution \( P_u(\zeta) \) approaches

\[ \prod_u \left[ P(u|\zeta)^{N P(u|\zeta_{\text{true}})} \right] \exp \left( N \sum_u P(u|\zeta_{\text{true}}) \ln P(u|\zeta) \right) \]

which exhibits a sharp peak at \( \zeta = \zeta_{\text{true}} \). For large \( N \), \( P_u(\zeta) \) is nonzero only in the vicinity of \( \zeta_{\text{true}} \). This justifies a Taylor expansion around \( \zeta_{\text{true}} \), leading to the Gaussian form \( P_u(\zeta) \sim e^{-[\zeta - \zeta_u]^2/(2 \sigma^2)} \) with a standard deviation \( \sigma \equiv 1/\sqrt{N F(\zeta_{\text{true}})} \). Then we have \( \zeta_{\text{M}} = \zeta_{\text{B}} = \zeta_{\text{true}} \) and \( \delta \zeta_{\text{M}} = \delta \zeta_{\text{B}} = \sigma \), so both estimators are unbiased and optimal in the large \( N \) limit.

Usually, calculating \( \zeta_{\text{M}} \) [Eq. (C1)] or \( \zeta_{\text{B}} \) [Eq. (C2)] requires intensive computational costs. The situation simplifies when the data comes from binary-outcome measurements, i.e., measurements that yield only two possible outcomes (denoted by \(+\) and \(-\)) according to the probability distribution \( P(\pm|\zeta) \). In this case, let \( N_+ \) (\( N_- \)) denote the number of outcome \(+\) (outcome \(-\)) from the \( N \) measurements, then solving \( P(+/\zeta) - P(-/\zeta) = (N_+ - N_-)/N \) for \( \zeta \) gives a simple estimator that is unbiased and optimal for large \( N \) [98]. In this work, we always adopt the maximum likelihood estimator.

**Appendix D: An example**

Here we follow the three standard steps outlined in Fig. 1 to estimate the level splitting \( \omega \) of a spin-1/2 Hamiltonian

\[ \hat{H}_0 = \frac{1}{2} \omega \hat{\sigma}_z \]

by monitoring its free precession. For convenience we define \( \mathbf{e}_\phi \) as a unit vector in the \( xy \) plane with azimuth \( \phi \).

For step 1, we assume the initial state to be \( |\psi_{\text{in}}\rangle = \cos(\Theta/2)|\uparrow\rangle + e^{i\phi} \sin(\Theta/2)|\downarrow\rangle \), where the controlling parameters \( \Theta \) and \( \Phi \) are to be optimized. Next, the spin-1/2 undergoes \( \omega \)-dependent free precession for an interval \( \tau \) into the final state \( |\psi_{\text{fin}}\rangle = e^{-i\omega \tau \sigma_z/2} |\psi_{\text{in}}\rangle \). The QFI in the final state is calculated by Eq. (B2) as

\[ F = \tau^2 \sin^2 \Theta \]

which is independent of \( \Phi \). To maximize \( F \), we set \( \Theta = \pi/2 \), and leave \( \Phi \) arbitrary, i.e., any initial state whose average spin lies in the \( xy \) plane is optimal. For specificity, we set \( \Phi = 0 \), so the initial state is the \( \hat{\sigma}_z = +1 \) eigenstate and the final state is the \( \hat{\sigma}_+ \cdot \mathbf{e}_\omega = +1 \) eigenstate, as shown in Fig. 7.

**FIG. 7.** (Color online). Free precession of a spin-1/2 around the \( z \) axis by an angle \( \omega \tau \). The red arrows denote the initial and final spin orientation and the blue arrow denotes the optimal measurement axis determined by the symmetric logarithmic derivative operator \( \hat{L}_{\omega\text{true}} \).

Interestingly, the QFI can be increased indefinitely by increasing the evolution time \( \tau \), indicating the time as a valuable quantum resource.

For step 2, we need to find optimal measurements to convert all the QFI into the CFI. There are two ways to find optimal measurements. The first one is to use the general conclusion [70] that the projective measurement on the symmetric logarithmic derivative operator \( \hat{L}_{\omega\text{true}} \) is optimal (see Appendix A). Since the final state is pure, we have

\[ \hat{L}_{\omega} = 2 \delta_{\omega} (|\psi_{\omega}\rangle\langle\psi_{\omega}|) = \tau \hat{\sigma}_- + t_{\omega/2+\omega\text{true}} \]

i.e., measuring \( \hat{\sigma}_- + t_{\omega/2+\omega\text{true}} \) (blue arrow in Fig. 7) is optimal. Actually, this measurement gives two possible outcomes \( \pm 1 \) according to the distribution \( P(\pm 1|\omega) = [1 \pm \sin(\omega - \omega_{\text{true}})|\tau]/2 \) and the CFI is computed from Eq. (B1) as \( F(\omega)|_{\omega = \omega_{\text{true}}} = \tau^2 = F \). Physically, this amounts to measuring the spin-1/2 along an axis (blue arrow in Fig. 7) perpendicular to the spin orientation of the final state \( |\psi_{\omega}\rangle \) (red arrow in Fig. 7) to maximize the dependence of the measurement distribution on the parameter \( \omega \). However, since \( \omega_{\text{true}} \) is unknown, complicated adaptive measurements are necessary. The second method is to consider a projective measurement on the spin-1/2 along a general axis parametrized by polar angle \( \theta \) and azimuth \( \varphi \), which gives the measurement distribution \( P(\pm 1|\omega) = [1 \pm \sin \theta \cos (\omega \tau - \varphi)]/2 \) and hence the CFI

\[ F = \tau^2 \frac{\sin^2 \theta \sin^2 (\omega \tau - \varphi)}{1 - \sin^2 \theta \cos^2 (\omega \tau - \varphi)} \]
To maximize $F$, we set $\theta = \pi/2$, then $F = \tau^2$ attains the QFI, i.e., measuring the spin-1/2 along an arbitrary axis in the $xy$ plane form a family of optimal measurements. For specificity we set $\varphi = 0$, corresponding to measuring $\delta \zeta$.

For step 3, suppose we have no prior knowledge about $\omega$ before the measurements. Next we repeat the initialization-evolution-measurement cycle twice and obtain two outcomes $u \equiv (u_1, u_2)$. Upon getting these outcomes, our knowledge about $\omega$ is immediately refined to (up to a constant normalization factor) $P_\omega(\omega) \sim P(u_1|\omega)P(u_2|\omega)$, e.g., if both outcomes are $+1$, then $P_\omega(\omega) \sim \cos^2(\omega \tau/2)$ shows many peaks at integer multiples of $2\pi/\tau$, corresponding to an infinite number of maximum likelihood estimators $\omega_M = 2n\pi/\tau (n \in \mathbb{Z})$. If both outcomes are $-1$, then $P_\omega(\omega) \sim \sin^2(\omega \tau/2)$ shows maxima at odd multiples of $\pi/\tau$, corresponding to $\omega_M = (2n + 1)\pi/2\tau (n \in \mathbb{Z})$. If one outcome is $+1$ and the other is $-1$, then $P_\omega(\omega) \sim \sin(\omega \tau)$ and $\omega_M = (2n + 1)\pi/2\tau (n \in \mathbb{Z})$. In any case, the maximum likelihood estimator is not unique, because the measurement distribution $P(u|\omega)$ is an even function of $\omega$ with a period $2\pi/\tau$, so $\omega$ and $-\omega$ (or $\omega$ and $\omega + 2\pi/\tau$) give exactly the same measurement distribution and hence cannot be distinguished. Such ambiguity can be eliminated by combining the information gained from measurements with different evolution time $\tau$ \[63, 64\].

**Appendix E: Feedback: adaptive measurement protocols**

As mentioned before, usually the CFI depends on the parameter $\zeta$ to be estimated, so optimizing the initial state, the evolution process, and the measurement scheme requires knowledge about $\zeta$, which is unknown. A standard solution is adaptive measurement protocols: after each initialization-evolution-measurement cycle, the measurement outcome is immediately used to refine our knowledge about $\zeta$, which in turn is used to optimize the next cycle, as shown in Fig. 1.

There are two categories of adaptive protocols. The first category focuses on maximizing the CFI \[83, 84\]. Suppose the CFI $F(\zeta, \theta)$ depends on $\zeta$ and some parameters $\theta$ that control the initialization, evolution, and measurement processes. The simplest idea is to tune $\theta$ to maximize the $F(\zeta, \theta)$. However, usually the optimal $\theta$ leading to maximal CFI depends on the unknown parameter $\zeta$. Suppose at the end of the $(n-1)$th initialization-evolution-measurement cycle, our knowledge is quantified by a distribution $P(\zeta)$, then a natural solution is to choose $\theta$ in the $n$th cycle to maximize the CFI averaged over the distribution of $\zeta$, i.e.,

$$F(\theta) \equiv \int F(\zeta, \theta)P(\zeta)d\zeta \approx F(\zeta_M, \theta),$$

where the second step is valid when the maximum of $P(\zeta)$ at

$$\zeta_M \equiv \arg \max P(\zeta)$$

is very sharp compared with $F(\zeta, \theta)$.

The second category focuses on optimizing the expected information gain from the estimator \[63, 64, 85, 86\]. At the end of the $(n-1)$th cycle, our knowledge about $\zeta$ is quantified by the distribution $P(\zeta)$. In the $n$th cycle with the controlling parameters $\theta$, the measurement distribution is $P_\theta(u|\zeta)$, which depend on $\zeta$ and $\theta$. If the measurement outcome of this cycle is $u$, then our knowledge about $\zeta$ would be updated to the distribution

$$P_{u,\theta}(\zeta) \sim P(\zeta)P_\theta(u|\zeta),$$

the maximum likelihood estimator would be $\zeta_M(u, \theta) \equiv \arg \max P_{u,\theta}(\zeta)$, and its uncertainty $\delta \zeta(u, \theta)$ would be given by Eq. (E3) with $P_\theta(\zeta) \rightarrow P_{u,\theta}(\zeta)$ and $\zeta_M(u) \rightarrow \zeta_M(u, \theta)$. Since the probability for this outcome $u$ to occur is given by $P_\theta(u|\zeta)$ averaged over the distribution $P(\zeta)$, i.e., $P_\theta(u) \equiv \int P_\theta(u|\zeta)P(\zeta)d\zeta$, we should choose $\theta$ in the $n$th cycle to minimize the expected uncertainty:

$$\overline{\delta \zeta}(\theta) \equiv \sum_u P_\theta(u)\delta \zeta(u, \theta).$$

When the maximum of $P(\zeta)$ at $\zeta_M$ [Eq. (E2)] is very sharp, we have $P_\theta(u) \approx P_\theta(u|\zeta_M)$, so

$$\overline{\delta \zeta}(\theta) \approx \sum_u P_\theta(u|\zeta_M)\delta \zeta(u, \theta).$$

The key idea of this adaptive scheme is to optimize the controlling parameters of the next cycle to minimize the expected uncertainty at the end of that cycle [Eq. (E4) or (E5)], so it is known as the locally optimal adaptive scheme \[85, 86\]. A straightforward extension is to optimize simultaneously the $M$ controlling parameters of the next $M$ cycles to minimize the expected uncertainty at the end of these cycles. Increasing $M$ improves the performance at the cost of exponentially increasing computational cost \[59\], so the $M = 1$ scheme is the most widely used one in Hamiltonian parameter estimation \[63, 64, 85, 86\].

Although both the CFI-based adaptive scheme and the locally optimal adaptive scheme have been widely used, their connection remains unclear. Here we prove their equivalence in the limit of very accurate knowledge $P(\zeta)$ at the beginning of the $n$th initialization-evolution-measurement cycle, as quantified by a sharp Gaussian distribution $P(\zeta) = e^{-(\zeta - \zeta_M)^2/(2\sigma^2)}/(\sqrt{2\pi}\sigma)$ with a small standard deviation $\sigma$. This justifies a Taylor expansion of Eq. (E3) around $\zeta_M$, which gives $P_{u,\theta}(\zeta) \sim e^{-(\zeta - \zeta_M)^2/(2\sigma_n^2)}$ with

$$\sigma_n^2 = \sigma^2 - \left(\frac{\partial^2 \ln P_\theta(u|\zeta)}{\partial \zeta^2}\Big|_{\zeta_M}\right) \Rightarrow \sigma_n^2 = \sigma + \frac{\sigma^3}{2} \left(\frac{\partial^2 \ln P_\theta(u|\zeta)}{\partial \zeta^2}\right)_{\zeta_M}.$$  

Then we have $\zeta_M(u, \theta) = \zeta_M(\theta)$ and $\delta \zeta(u, \theta) = \sigma_n$, so the expected uncertainty in Eq. (E5) becomes

$$\overline{\delta \zeta}(\theta) \approx \sigma - \frac{1}{2} \sigma^3 F(\zeta_M, \theta),$$

thus minimizing $\overline{\delta \zeta}(\theta)$ amounts to maximizing $F(\zeta_M, \theta)$ in Eq. (E1).
Note1, in principle, we can also apply time-dependent control on the spin during the evolution to engineer the final state and then maximize $\mathcal{F}_w$ or $\mathcal{F}_f$ in the final state by optimizing this control. However, such optimization for a time-dependent, open quantum system is still an open issue [84, 100, 101].