Saturating quantum advantages in postselected metrology with the positive Kirkwood-Dirac distribution

Sourav Das [1,2,*] Subhrajit Modak [1,3,†] and Manabendra Nath Bera [1,‡]

[1]*Department of Physical Sciences, Indian Institute of Science Education and Research (IISER), Mohali, Punjab 140306, India*

[2]*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*

[3]*Department of Physics, Indian Institute of Technology Madras, Chennai 600036, India*

(Received 14 September 2021; revised 18 September 2022; accepted 20 March 2023; published 11 April 2023)

Weak amplification and other postselection-based metrological protocols can enhance precision while estimating small parameters, outperforming postselection-free protocols. In general, these enhancements are largely constrained because the protocols yielding higher precision are rarely obtained due to a lower probability of successful postselection. It is shown that this precision can further be improved with the help of quantum resources like entanglement and negativity in the quasiprobability distribution. However, these quantum advantages in attaining considerable success probability with large precision are bounded irrespective of any accessible quantum resources. The advantage is being considered only within the scope of postselected metrology. Here we derive a bound of these advantages in postselected metrology, establishing a connection with weak value optimization where the latter can be understood in terms of the geometric phase. We introduce a scheme that saturates the bound, yielding anomalously large precision. Moreover, we prove that these advantages can be achieved with positive quasiprobability distribution. We also provide an optimal metrological scheme using a three-level nondegenerate quantum system.

DOI: 10.1103/PhysRevA.107.042413

I. INTRODUCTION

Parameter estimation, which is central to mathematical statistics, is an elementary problem in information theory. Its main objective is to construct and evaluate various methods that can estimate the values of parameters of either an information source or a communication channel. In estimation theory, the Cramér-Rao inequality [1,2] expresses a lower bound on the variance of unbiased estimators θ_e stating that the variance of any such estimator $\mathrm{Var}(\theta_e)$ is at least as high as the inverse of the Fisher information, $\mathcal{F}(\theta)$:

$$\mathrm{Var}(\theta_e) \geqslant \frac{1}{\mathcal{F}(\theta)}.$$

The Fisher information quantifies the amount of information a sample conveys about an unknown parameter. In other words, it tells us how well one can measure a parameter, given a certain amount of data. A common metrological task is concerned with designing an optimized experimental setup that minimizes the estimator's error by maximizing the Fisher information [3].

It has been shown that utilization of quantum resources can improve the precision of parameter estimations beyond classical limits. This idea is at the basis of the continuously growing research area of quantum metrology that aims at reaching the fundamental bounds in metrology by exploiting quantum probes. The central quantity in quantum metrology is the quantum Fisher information (QFI), which is the optimal Fisher information of a metrological protocol over different measurement settings [4–7]. When N number of classical probes, each interacting onc at a time with the system under study, the variance of the estimator optimally scales as $N^{-\frac{1}{2}}$, and it stands for the probes that are at most classically correlated. This limit is called the standard quantum limit (SQL), and it stands for the probes that are at most classically correlated. Using quantum probes, this scale can be optimally enhanced to N^{-1}, called the Heisenberg limit (HL) [8–10]. The quest for measurement schemes surpassing the SQL has inspired a variety of clever strategies, employing squeezing of the vacuum [11–15], optimizing the probing time [16], monitoring the environment [17,18], and exploiting non-Markovian effects [19,20]. Besides the fundamental interest about ultimate precision limits, quantum metrology presents different applications, such as measurement on biological systems [21,22], gravitational wave detection [23], atomic clocks [24–26], interferometry with atomic and molecular matter waves [27–29], Hamiltonian estimation [30–33], and other general sensing technologies [34,35].

It is considered that the HL represents a fundamental limit on the sensitivity of quantum measurements. However, different studies have shown that interactions among particles may be a valuable resource for quantum metrology, allowing scaling beyond the HL [36–38]. Naturally, the question arises: What are the alternative protocols that can be relevant to achieving such scaling? A recent theoretical study has reported that the postselected quantum experiments can be used to overcome the HL by enabling a quantum state

*sourav.iisermohali@gmail.com
†modoksuvrojit@gmail.com
‡mnbera@gmail.com

to carry more Fisher information [39]. The reason behind this benefit is claimed to be the negative quasiprobability distribution [40–46] which is an important manifestation of nonclassicality. Also, it has been shown that improved advantages can be attained when properly conditioned experiments are performed. However, this advantage comes with a lower rate of successful postselection [47,48].

We ask, in this article, if this advantage be bounded fundamentally irrespective of any accessible quantum resources. We conclude that it can. Using geometric arguments, we derive a bound of these advantages using weak value optimization and show that the intrinsic weak values of the system observable play a key role in bounding this advantage. We construct a preparation-and-postselection procedure that saturates the bound using a three-level nondegenerate quantum system. Surprisingly, at the saturation point, the quasiprobability distribution of this setting turns out to be positive. So far, the negative quasiprobability distribution is considered essential for postselected QFI to overcome the HL. Our scheme achieves the same without accounting for any negative or nonreal elements in the quasiprobability distribution. Finally, we propose an alternative way to understand this quantum advantage.

II. POSTSELECTED QUANTUM FISHER INFORMATION

A typical situation in quantum parameter estimation is to estimate a parameter θ that is encoded on the system state through some general quantum evolution $\hat{\rho}_\theta = \mathcal{E}_\theta(\hat{\rho})$, where $\mathcal{E}_\theta$ is a trace-preserving completely positive map. The parameters could be the phases of light in interferometers, unitary phase shifts, the decay constant of an atom, the strength of magnetic or gravitational fields, etc. The efficiency of an estimation procedure is characterized in terms of the mean-square estimation error (MSE),

$$\mathrm{MSE}(\theta_{\mathrm{true}}) = \mathbb{E}_{\mathrm{data}}\{[\theta(\mathrm{data}) - \theta_{\mathrm{true}}]^2\}, \tag{1}$$

where θ is an estimator for θ_{true} and $\mathbb{E}_{\mathrm{data}}\{.\}$ denotes an expectation over data. For unbiased estimators, a strict lower bound on the MSE—the *quantum Cramér-Rao bound* (QCRB)—is expressed in terms of the quantum Fisher information $\mathcal{F}_Q(\theta)$ associated with the state $\hat{\rho}_\theta$ that encodes the parameter. Under measurements on N copies of the system, the QCRB is expressed as

$$\mathrm{MSE}(\theta) \geqslant \frac{1}{N\mathcal{F}_Q(\theta)}. \tag{2}$$

Consider a quantum experiment that outputs the state $\hat{\rho}_\theta = \hat{U}(\theta)\hat{\rho}_0\hat{U}^\dagger(\theta)$, where $\hat{\rho}_0$ undergoes the unitary transformation $\hat{U}(\theta) = e^{-i\hat{A}\theta}$ driven by $\hat{A}$ with $\theta \in \mathbb{R}$. If $\hat{\rho}_\theta$ is pure, such that $\hat{\rho}_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, the optimized quantum Fisher information can be written as

$$\mathcal{F}_Q(\theta|\hat{\rho}_\theta) = 4\mathrm{Var}(\hat{A})_{\hat{\rho}_0}. \tag{3}$$

The Fisher information can be further optimized over all input states, which gives

$$\max_{\hat{\rho}_0}\{\mathcal{F}_Q(\theta|\hat{\rho}_\theta)\} = 4\max_{\hat{\rho}_0}\{\mathrm{Var}(\hat{A})_{\hat{\rho}_0}\} = (\Delta a)^2, \tag{4}$$

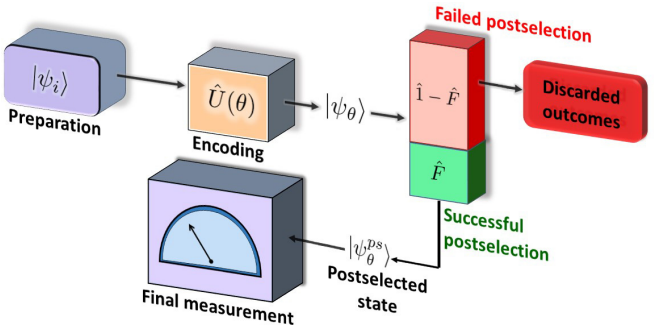

FIG. 1. A scheme for postselected metrology First, an input quantum state $|\psi_i\rangle$ undergoes a unitary transformation $\hat{U}(\theta) = e^{-i\theta\hat{A}}$: $|\psi_i\rangle \to |\psi_\theta\rangle$. Second, the quantum state is subjected to a postselective measurement with the projectors $\{\hat{F}, \hat{1} - \hat{F}\}$. After the successful postselection with $\hat{F}$, the updated (renormalized) state becomes $|\psi_\theta^{ps}\rangle = \hat{F}|\psi_\theta\rangle/(\langle\psi_\theta|\hat{F}|\psi_\theta\rangle)^{1/2}$, which is further analyzed to estimate the parameters.

where Δa is the difference between the maximum and the minimum eigenvalues of $\hat{A}$. In order to show how postselection could help retrieve more information per measurement, we present a short review of the postselected prepare-measure experiment. According to the protocol, as shown in Fig. 1, a selective measurement takes place after $\hat{U}(\theta)$ but before the final measurement. The renormalized quantum state that passes the postselection is $|\psi_\theta^{\mathrm{ps}}\rangle \equiv |\Psi_\theta^{\mathrm{ps}}\rangle/\sqrt{p_\theta^{\mathrm{ps}}}$, where the unnormalized state is defined as $|\Psi_\theta^{\mathrm{ps}}\rangle \equiv \hat{F}|\psi_\theta\rangle$, with $\hat{F} = \sum_{f\in\mathcal{F}^{\mathrm{ps}}}|f\rangle\langle f|$ being the postselecting projection operator, where $\mathcal{F}^{\mathrm{ps}}$ is the chosen set of postselection and $p_\theta^{\mathrm{ps}} \equiv \mathrm{Tr}(\hat{F}\hat{\rho}_\theta)$ is the probability of postselection. Finally, the postselected state goes through an optimized measurement protocol. Compiling all the steps leads to [39]

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\mathrm{ps}}\big) = 4\big\langle\dot{\Psi}_\theta^{\mathrm{ps}}|\dot{\Psi}_\theta^{\mathrm{ps}}\big\rangle\frac{1}{p_\theta^{\mathrm{ps}}} - 4\big|\big\langle\dot{\Psi}_\theta^{\mathrm{ps}}|\Psi_\theta^{\mathrm{ps}}\big\rangle\big|^2\frac{1}{\big(p_\theta^{\mathrm{ps}}\big)^2}, \tag{5}$$

where $|\dot{\Psi}_\theta^{\mathrm{ps}}\rangle \equiv \partial_\theta|\Psi_\theta^{\mathrm{ps}}\rangle$.

This gives the quantum Fisher information available from a quantum state after its postselection. Evidently, $\mathcal{F}_Q(\theta|\psi_\theta^{\mathrm{ps}})$ exceeds $\mathcal{F}_Q(\theta|\hat{\rho}_\theta)$, since $p_\theta^{\mathrm{ps}} \leqslant 1$. Moreover, if the postselections result in a quasiprobability distribution (i.e., with negative entries), the rate of improvement in retrieving information per measurement can even surpass the standard limit of the Fisher information [39]. The postselected quantum Fisher information can be re-expressed in terms of the quasiprobability distribution:

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\mathrm{ps}}\big) = 4\sum_{\substack{a,a',\\ f\in\mathcal{F}^{\mathrm{ps}}}}\frac{\mathcal{K}^{\hat{\rho}_\theta}_{a,a',f}}{p_\theta^{\mathrm{ps}}}aa' - 4\left|\sum_{\substack{a,a',\\ f\in\mathcal{F}^{\mathrm{ps}}}}\frac{\mathcal{K}^{\hat{\rho}_\theta}_{a,a',f}}{p_\theta^{\mathrm{ps}}}a\right|^2, \tag{6}$$

where $\mathcal{K}^{\hat{\rho}_\theta}_{a,a',f} = \langle f|a\rangle\langle a|\hat{\rho}_\theta|a'\rangle\langle a'|f\rangle$ refers to an element of the doubly extended Kirkwood-Dirac (KD) quasiprobability distribution [40,41], defined in terms of eigenbases of $\hat{A}$ and $\hat{F}$. This is an extension of the standard KD distribution which is obtained under two weak measurements followed by a strong measurement performed sequentially on the

system. If $\hat{A}$ commutes with $\hat{F}$, as they do classically, then they share an eigenbasis for which $\mathcal{K}^{\hat{\rho}_\theta}_{a,a',f}/p^{\rm ps}_\theta \in [0, 1]$, and the postselected quantum Fisher information is bounded as $\mathcal{F}_Q(\theta|\psi^{\rm ps}_\theta) \leqslant (\Delta a)^2$. In contrast to this, if the quasiprobability distribution contains negative values, the postselected quantum Fisher information can be seen to violate the standard bound: $\mathcal{F}_Q(\theta|\psi^{\rm ps}_\theta) > (\Delta a)^2$. We will see shortly how these nonclassical advantages can be accommodated in any quantum experiments to achieve anomalously large quantum Fisher information. In fact, this gain in information can be related to the weak values of the system observable. For that, first we rewrite Eq. (6) in the operator form:

$$\mathcal{F}_Q\big(\theta|\psi^{\rm ps}_\theta\big) = \frac{4}{p^{\rm ps}_\theta}{\rm Tr}[\hat{F}\hat{A}\hat{U}(\theta)\hat{\rho}_0\hat{U}(\theta)^\dagger\hat{A}] - \frac{4}{\big(p^{\rm ps}_\theta\big)^2}|{\rm Tr}[\hat{F}\hat{U}(\theta)\hat{\rho}_0\hat{U}(\theta)^\dagger\hat{A}]|^2. \quad (7)$$

The aim is to choose $\hat{F}$ and $\hat{\rho}_0$ in a way that $\mathcal{F}_Q(\theta|\psi^{\rm ps}_\theta)$ approaches the maximum. We begin with a pure initial state $|\psi_i\rangle$ and the postselction is carried out with the projector $\hat{F} = \sum_{f_k\in\mathcal{F}^{\rm ps}}|f_k\rangle\langle f_k|$ on the updated state $|\psi_\theta\rangle = \hat{U}(\theta)|\psi_i\rangle$, where $\mathcal{F}^{\rm ps}$ is the postselection basis. Then the postselection probability becomes

$$p^{\rm ps}_\theta = \sum_{f_k\in\mathcal{F}^{\rm ps}}|\langle\psi_\theta|f_k\rangle|^2, \quad (8)$$

and Eq. (7) leads to

$$\mathcal{F}_Q\big(\theta|\psi^{\rm ps}_\theta\big)\big(p^{\rm ps}_\theta\big)^2 = 4\sum_{i<j}p^{\rm ps}_{\theta i}p^{\rm ps}_{\theta j}\big|A^i_w - A^j_w\big|^2, \quad (9)$$

where $A^k_w = \frac{\langle\psi_\theta|\hat{A}|f_k\rangle}{\langle\psi_\theta|f_k\rangle}$ and $p^{\rm ps}_{\theta k} = |\langle\psi_\theta|f_k\rangle|^2$. It can be seen that there is no enhancement in postselected QFI when the system is postselected with a rank-1 projector.

Note, A^k_w is nothing but the weak values corresponding to the observable $\hat{A}$ between the preselected state $|\psi_\theta\rangle$ and the kth component of the postselected state $|f_k\rangle$, and $p^{\rm ps}_{\theta k}$ is the postselection probability [49] (see the Appendices for details). Therefore, optimizing enhancement in postselected metrology can be connected to the weak value optimization, as we discuss below.

III. WEAK VALUE OPTIMIZATION: A GEOMETRIC INTERPRETATION

In general, a large weak value appears when it is less likely to have successful postselection of the system. It indicates that large weak values are obtained but very rarely. It has been shown that by exploiting quantum resources, e.g., entanglement, squeezed states, etc., the success probability can be improved for a fixed weak value. However, the advantage in attaining considerable success probability with a large weak value is bounded irrespective of any accessible quantum resources. To optimize the advantage, one can either start by fixing the weak value or keeping the success probability fixed. Usually, these are attained from separate optimization protocols. In order to access these advantages simultaneously from a single protocol, we can start with optimizing a combined quantity η:

$$\eta(A_w, p_s) = p_s|A_w|^2. \quad (10)$$

We refer to this quantity as the *efficiency* of a weak value metrological protocol. To check how efficient a protocol is, one needs to quantify the gain in terms of $|A_w|$ along with a cost $\frac{1}{p_s}$ to access the same. Evidently, minimal cost implies higher probability of successful postselection. The efficiency is bounded by the following relation:

$$\eta(A_w, p_s) \leqslant ||\hat{A}^2||_{\rm op}, \quad (11)$$

where the operator norm is defined as $||\hat{X}||_{\rm op} := \sup_{|\phi\rangle\in\mathcal{H}}\{\langle\phi|\hat{X}|\phi\rangle : \langle\phi|\phi\rangle = 1\}$. Once the upper bound is known, the optimization procedure can be initiated using a proper trial function. To make it more insightful, we would like to explore the geometric connection behind the process of optimization. A detailed analysis will soon reveal how the geometric phase appears in the context of efficiency of a metrological experiment. To establish the connection, we start by taking the observable $\hat{A} = \sum_k a_k|a_k\rangle\langle a_k|$. Substituting it into Eq. (10) leads to

$$\eta(A_w, p_s) = \Big|\sum_k a_k|\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle| \times \exp\big[i\Phi^{a_k}_g(|\psi_i\rangle, |a_k\rangle, |\psi_f\rangle)\big]\Big|^2, \quad (12)$$

where $\Phi^{a_k}_g(|\psi_i\rangle, |a_k\rangle, |\psi_f\rangle) := \arg(\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle\langle\psi_i|\psi_f\rangle)$ is the Bergman angle, widely known as the geometric phase [50]. In the realm of quantum states when cyclic transition occurs starting from a preselected state $|\psi_i\rangle$ to the same state via a path that connects $|a_k\rangle$ and $|\psi_f\rangle$ through geodesic lines on the Bloch sphere, the final state acquires an excess phase over $|\psi_i\rangle$. This phase is proportional to the solid angle at the center, subtended by the geodesic triangle with vertices at $|\psi_i\rangle$, $|\psi_f\rangle$, and $|a_k\rangle$. Now, we aim to optimize the protocol for maximum efficiency. One way to ensure this is to keep the function under summation positive for all k. Any exception to this will not lead to the desired optimization. To saturate the bound, the postselected state is taken to be parallel to $\hat{A}|\psi_i\rangle$, i.e.,

$$|\psi_f\rangle = \frac{\hat{A}|\psi_i\rangle}{\sqrt{\langle\psi_i|\hat{A}^2|\psi_i\rangle}}, \quad (13)$$

which is similar to the case considered in Ref. [51]. Evaluating the expression for the Bergman angle, we obtain

$$\Phi^{a_k}_g(|\psi_i\rangle, |a_k\rangle, |\psi_f\rangle) = \arg\left[a_k\frac{|\langle\psi_i|a_k\rangle|^2}{\sqrt{\langle\psi_i|\hat{A}^2|\psi_i\rangle}}\right] + \arg(\langle\psi_f|\psi_i\rangle). \quad (14)$$

Note that, instead of a_k, all the terms inside the first argument function are positive, which simplifies the expression:

$$\Phi^{a_k}_g(|\psi_i\rangle, |a_k\rangle, |\psi_f\rangle) = \arg[{\rm sgn}(a_k)] + \arg(\langle\psi_f|\psi_i\rangle). \quad (15)$$

This straightforwardly leads to

$$\exp\big[i\Phi^{a_k}_g(|\psi_i\rangle, |a_k\rangle, |\psi_f\rangle)\big] = {\rm sgn}(a_k)\exp(i\phi), \quad (16)$$

where $\mathrm{sgn}(a_k) = \frac{|a_k|}{a_k}$ is the sign function and $\phi = \arg(\langle\psi_f|\psi_i\rangle)$ is a constant phase which does not contribute to the efficiency. The maximum efficiency is attained when $\Phi_g^{a_k}$ depends on the eigenvalues a_k via a sign function only. This ensures that the terms under summation in Eq. (12) are positive for all k, leading to maximum efficiency. An example using the spin-1/2 system is outlined in the Appendices. It is worth noting that the optimization in terms of efficiency, following the geometric argument, is more general in the sense that it takes into account both the weak values and the postselection probability.

IV. BOUNDING POSTSELECTED METROLOGY THROUGH WEAK VALUE OPTIMIZATION

In the same line, as mentioned in the context of efficiency of a weak value amplification, we can also assign a trade-off relation between the probability of postselection and the information obtained upon postselection. We begin with defining a similar quantity as metrological efficiency here as well.

Definition 1. For any arbitrary state preparation and postselection, the efficiency of a postselected metrological protocol ξ^{ps} with the postselected Fisher information $\mathcal{F}_Q$ and the total probability of successful postselection p_θ^{ps} is defined as

$$\xi^{\mathrm{ps}}\left(p_\theta^{\mathrm{ps}};\mathcal{F}_Q\right) = p_\theta^{\mathrm{ps}}\mathcal{F}_Q. \tag{17}$$

This efficiency cannot be arbitrarily large, as mentioned in the theorem below.

Theorem 2. The efficiency of a protocol for postselected metrology is bounded according to the following inequality:

$$\xi^{\mathrm{ps}}\left(p_\theta^{\mathrm{ps}};\mathcal{F}_Q\right) \leqslant 4||\hat{A}^2||_{\mathrm{op}}. \tag{18}$$

The highest efficiency is achieved under the following condition,

$$\langle\psi_i|\hat{A}|\psi_i\rangle = 0 \ \text{ and } \ \forall f_k \notin \mathcal{F}^{\mathrm{ps}}, \ A_w^k = 0. \tag{19}$$

Proof. Since the Fisher information decreases when the states are mixed, we can expect that the maximum efficiency will be achieved for pure states only. Here, we start with preparing the system in a pure initial state $|\psi_i\rangle$ which evolves to $|\psi_\theta\rangle$ after the parameter θ is encoded via unitary $\exp(-i\hat{A}\theta)$. Then, the efficiency of the protocol can be expressed in terms of the standard KD distribution and efficiency in weak value amplification as

$$\xi^{\mathrm{ps}}\left(p_\theta^{\mathrm{ps}};\mathcal{F}_Q\right) = 4\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}}\eta\left(p_{\theta k}^{\mathrm{ps}}, A_w^k\right) - \frac{4}{p_\theta^{\mathrm{ps}}}\left|\sum_m\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m\mathcal{K}_{a_m,f_k}^{|\psi_\theta\rangle}\right|^2, \tag{20}$$

where $\mathcal{K}_{a_m,f_k}^{|\psi_\theta\rangle} := \langle\psi_\theta|a_m\rangle\langle a_m|f_k\rangle\langle f_k|\psi_\theta\rangle$ is the standard KD distribution (see the Appendices for derivation). To maximize the efficiency we treat the first and second terms in the right-hand side of this expression independently. Applying Bessel's inequality for the first term, we obtain

$$\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}}\eta\left(p_{\theta k}^{\mathrm{ps}}, A_w^k\right) = \sum_{f_k\in\mathcal{F}^{\mathrm{ps}}}|\langle\psi_\theta|\hat{A}|f_k\rangle|^2 \leqslant ||\hat{A}^2||_{\mathrm{op}}. \tag{21}$$

Since the second term is a non-negative quantity, the inequality (18) is always respected. ∎

Now we investigate the condition to achieve optimal efficiency. The inequality (21) saturates under the following condition,

$$\sum_{f_k\notin\mathcal{F}^{\mathrm{ps}}}\eta\left(p_{\theta k}^{\mathrm{ps}}, A_w^k\right) = 0. \tag{22}$$

Since all the terms inside the summation sign are positive, the condition is equivalent with $\forall f_k \notin \mathcal{F}^{\mathrm{ps}}$, $\langle\psi_\theta|\hat{A}|f_k\rangle = 0$. This implies that all the intrinsic weak values with failed postselected states have to be zero separately. The condition for the second term in the right-hand side of Eq. (21) to be zero is $\sum_m\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m\mathcal{K}_{a_m,f_k}^{|\psi_\theta\rangle} = 0$, which implies $\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}}\langle\psi_\theta|\hat{A}|f_k\rangle\langle f_k|\psi_\theta\rangle = 0$. Putting together both the conditions, we obtain

$$\langle\psi_i|\hat{A}|\psi_i\rangle = 0 \ \text{ and } \ \forall f_k \notin \mathcal{F}^{\mathrm{ps}}, \ A_w^k = 0. \tag{23}$$

Under this condition, the inequality (18) saturates, which is termed as information-preserving postselected metrology.

Now, we point out a situation where we achieve the bound for a rank-2 postselection. Then, Eq. (9) reduces to $\mathcal{F}_Q(\theta|\psi_\theta^{\mathrm{ps}})(p_\theta^{\mathrm{ps}})^2 = 4p_{\theta1}^{\mathrm{ps}}p_{\theta2}^{\mathrm{ps}}|A_w^1 - A_w^2|^2$. First we set $A_w^2 = 0$. Under this condition, we get

$$\xi^{\mathrm{ps}}\left(p_\theta^{\mathrm{ps}};\mathcal{F}_Q\right) = 4\frac{p_{\theta1}^{\mathrm{ps}}p_{\theta2}^{\mathrm{ps}}}{p_{\theta1}^{\mathrm{ps}}+p_{\theta2}^{\mathrm{ps}}}|A_w^1|^2. \tag{24}$$

Imposing the weak value optimization (11) on the first postselection leads to

$$\xi^{\mathrm{ps}}\left(p_\theta^{\mathrm{ps}};\mathcal{F}_Q\right) = 4\frac{p_{\theta2}^{\mathrm{ps}}}{p_{\theta1}^{\mathrm{ps}}+p_{\theta2}^{\mathrm{ps}}}||\hat{A}^2||_{\mathrm{op}}. \tag{25}$$

Once $p_{\theta1}^{\mathrm{ps}}$ starts to tend towards zero, one can get closer to the saturation. Thus, we need to maintain $|f_2\rangle$ perpendicular to $\hat{A}|\psi_\theta\rangle$ and $|\psi_\theta\rangle$ parallel to $\hat{A}|f_1\rangle$, simultaneously. Also, we need to choose the initial state of the system in such a way that the average of $\hat{A}^2$ saturates under the optimality conditions. It is now evident from the geometric interpretation that the optimization for the first intrinsic weak value would be attained when the geometric phases between the preselection, the eigenstates of $\hat{A}$, and the postselection have certain dependence with the eigenvalues of $\hat{A}$. Below, we give an example to show how to achieve optimal quantum enhancement in the context of postselected metrology.

A. Information-preserving postselected metrology using a three-level quantum system

Consider the state space of the system is spanned by the basis set $\{|\lambda\rangle, |\tilde{\lambda}\rangle, |-\lambda\rangle\}$, which are the eigenvectors of the system observable $\hat{A}$ with corresponding eigenvalues λ, $\tilde{\lambda}$, and $-\lambda$, respectively. We want to estimate θ that lies close to its true value θ_0 and the difference between them is $\delta_\theta \equiv \theta - \theta_0$, with $|\delta_\theta| \ll 1$. First, we choose the postselection $\hat{F} = |f_1\rangle\langle f_1| + |f_2\rangle\langle f_2|$, with

$$|f_1\rangle = \frac{|\lambda\rangle - |-\lambda\rangle}{\sqrt{2}}, \tag{26}$$

$$|f_2\rangle = \frac{\cos\alpha|\lambda\rangle + \cos\alpha|-\lambda\rangle}{\sqrt{2}} + \sin\alpha|\tilde{\lambda}\rangle, \tag{27}$$

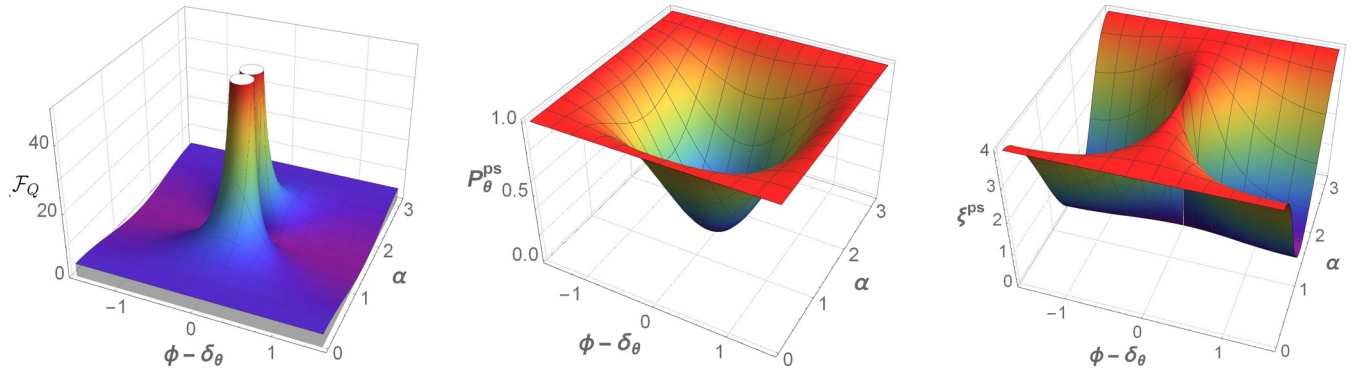

FIG. 2. Information-preserving postselected metrology. Figures from the left represent the postselected Fisher information $\mathcal{F}_Q$, the probability of successful postselection p_θ^{ps}, and the efficiency of the protocol, respectively, with different values of $\phi-\delta_\theta$ and α, for $\lambda=1$. The optimality condition is attained when $\phi\to\delta_\theta$. For more details, see the text.

where α is a real parameter. We also choose the initial state

$$|\psi_i\rangle=\frac{U^\dagger(\theta_0)[(\cos\phi+i\sin\phi)|\lambda\rangle+(\cos\phi-i\sin\phi)|-\lambda\rangle]}{\sqrt{2}}, \tag{28}$$

where the unitary $\hat{U}(\theta_0)=\exp(-i\hat{A}\theta_0)$. In this setting the evolved state becomes

$$|\psi_\theta\rangle=\hat{U}(\theta)|\psi_i\rangle=\frac{1}{\sqrt{2}}[(\cos\phi+i\sin\phi)e^{-i\delta_\theta\lambda}|\lambda\rangle + (\cos\phi-i\sin\phi)e^{i\delta_\theta\lambda}|-\lambda\rangle], \tag{29}$$

where ϕ is a real parameter. Putting this postselection and the initial state into Eq. (10), we find

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{ps}\big)p_\theta^{ps}=\frac{4\lambda^2\cos^2\alpha}{\sin^2(\phi-\lambda\delta_\theta)+\cos^2\alpha\cos^2(\phi-\lambda\delta_\theta)}. \tag{30}$$

The total postselection probability comes out of the following estimation:

$$p_\theta^{ps}=p_{\theta1}^{ps}+p_{\theta2}^{ps}=\sin^2(\phi-\lambda\delta_\theta)+\cos^2\alpha\cos^2(\phi-\lambda\delta_\theta), \tag{31}$$

along with the respective weak values

$$A_w^1=\frac{-i\lambda\cos(\phi-\lambda\delta_\theta)}{\sin(\phi-\lambda\delta_\theta)},\quad A_w^2=\frac{i\lambda\sin(\phi-\lambda\delta_\theta)}{\cos(\phi-\lambda\delta_\theta)}. \tag{32}$$

As ϕ approaches $\lambda\delta_\theta$, the condition for optimal quantum advantages gets satisfied. This can clearly be seen from the following expressions when compiling with appropriate limits:

$$\lim_{\phi\to\lambda\delta_\theta}p_\theta^{ps}=\cos^2\alpha, \tag{33}$$

$$\lim_{\phi\to\lambda\delta_\theta}\mathcal{F}_Q\big(\theta|\psi_\theta^{ps}\big)=4\lambda^2\sec^2\alpha,\quad\text{and} \tag{34}$$

$$\lim_{\phi\to\lambda\delta_\theta}\xi\big(p_\theta^{ps};\mathcal{F}_Q\big)=4\lambda^2. \tag{35}$$

Thus, in the aforementioned limit, we achieve the information-preserving protocol associated with the postselected metrology. Figure 2 reflects how the Fisher information, the probability of postselection, and the efficiency of the protocol are dependent on the change of parameters $\{\phi-\lambda\delta_\theta,\alpha\}$. The parameter α opens up a choice for the experimenter to decide the degree of precision without violating the restriction of information preservation. Before moving to further discussion, we should note that tuning ϕ in the limit $\phi\to\lambda\delta_\theta$ would not be an easy task. We must have a pre-estimated error range of the parameter θ beforehand; let us call it Δ_θ. The experimenter then needs to perform a trial run tuning ϕ with different values within the range $|\phi|\leqslant\lambda\Delta_\theta$. When Δ_θ is sufficiently small, reaching the optimal limit becomes easier. Even if the tuning is not perfect for reaching the optimal point, it gives us a range of different values of ϕ where we may experience anomalous QFI (see Fig. 2). To widen this range, one can consider increasing the value of λ.

In comparison to the setup [39], our protocol relaxes some experimental restrictions yet achieves anomalous QFI without any loss of information. Additionally, we can experience this anomalousness in this setup even without setting the $\delta_\theta\to0$ condition. Moreover, all these advantages can be obtained only by considering a three-level system. Thus, by construction, this setup is more profitable to attain the desired results.

B. Role of quasiprobalility in postselected metrology

In this section, we show that the protocols having positive KD distribution, which so far has been considered to be classical, can harness quantum advantages in postselected metrology.

We aim to prove this claim by the same example mentioned earlier. The KD distribution corresponding to the system state is positive everywhere when the optimality conditions are imposed (see Table I). However, under these conditions, we can obtain anomalously large QFI.

Here, we propose an alternative approach for pure states. We start by referring an identity that addresses the relation between the KD distribution and the classical joint probability distribution for two successive projective measurements [42]:

$$q_{a_m,f_k}^{\hat{\rho}}=\mathrm{Tr}(\hat{\rho}\hat{A}_m\hat{F}_k\hat{A}_m)+\frac{1}{2}\big\{\mathrm{Tr}[(\hat{\rho}-\hat{\rho}')\hat{F}_k] + \mathrm{Tr}\big[(\hat{\rho}-\hat{\rho}')\hat{F}_k^{\frac{\pi}{2}}\big]\big\}, \tag{36}$$

where $\hat{F}_k^{\frac{\pi}{2}} = \exp(-i\frac{\pi}{2}\hat{A}_m)\hat{F}_k\exp(i\frac{\pi}{2}\hat{A}_m)$ and $\hat{\rho}'$ is the post-measurement state achieved after a nonselective projective measurement on $\hat{\rho}$ with $\hat{A}_m$ and $(\hat{1} - \hat{A}_m)$. $\{\hat{A}_m\}$ and $\{\hat{F}_k\}$ are two different orthonormal sets of projectors. Here, the KD distribution is composed of two terms. The first term in the right-hand side is called the Wigner formula, which refers the classical joint probability distribution for a set of outcomes a_m and f_k upon a successive projective measurement on $\hat{\rho}$ first with $\hat{A}_m$ and then with $\hat{F}_k$. The second and third terms together can be considered as quantum modification terms which arise due to the the noncommutativity of measurement observables. Whenever the observables commute, this part vanishes. For simplicity, we denote the Wigner formula as $Q^{\hat{\rho}}_{a_m,f_k} := \mathrm{Tr}(\hat{\rho}\hat{A}_m\hat{F}_k\hat{A}_m)$.

Theorem 3. For pure states, the metrological efficiency $\xi^{\mathrm{ps}}(p^{\mathrm{ps}}_\theta;\mathcal{F}_Q) = 0$, when its KD distribution equals the corresponding Wigner formula,

$$\mathcal{K}^{\hat{\rho}_\theta}_{a_m,f_k} = Q^{\hat{\rho}_\theta}_{a_m,f_k}. \tag{37}$$

Proof. Upon solving to satisfy the above condition, one can straightforwardly obtain two conditions:

$$\langle\psi_\theta|a_m\rangle\langle a_m|f_k\rangle = 0 \tag{38}$$

or

$$\langle a_m|\psi_\theta\rangle\langle f_k|a_m\rangle = \langle f_k|\psi_\theta\rangle. \tag{39}$$

The condition (39) refers to the points where the KD distribution and the Wigner formula both take zero values. To keep the triviality aside, we neglect that and consider the condition (40). Recall Eq. (21),

$$\xi^{\mathrm{ps}}\big(p^{\mathrm{ps}}_\theta;\mathcal{F}_Q\big) = 4\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} p^{\mathrm{ps}}_{\theta k}|A^k_w|^2 - \frac{4}{p^{\mathrm{ps}}_\theta}\left|\sum_m\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m\mathcal{K}^{|\psi_\theta\rangle}_{a_m,f_k}\right|^2. \tag{40}$$

Decomposing the terms under summation and imposing the condition (40) lead to

$$\sum_{f\in\mathcal{F}^{\mathrm{ps}}} p^{\mathrm{ps}}_{\theta k}|A_k|^2 = p^{\mathrm{ps}}_\theta\left|\sum_m a_m\right|^2,$$

$$\sum_m\sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m\mathcal{K}^{|\psi_\theta\rangle}_{a_m,f_k} = \big(p^{\mathrm{ps}}_\theta\big)^2\left|\sum_m a_m\right|^2.$$

Replacing this to Eq. (40) yields

$$\xi^{\mathrm{ps}}\big(p^{\mathrm{ps}}_\theta;\mathcal{F}_Q\big) = 0. \tag{41}$$

See the Appendices for the detailed calculation. ∎

We have emphasized the point earlier that the KD distribution and the Wigner formula are not equal when the measurement observables do not commute. However, the KD distribution may still be positive in such cases. The example here also exhibits a similar feature (see Table II), where the Wigner formula differs from KD distribution. This observation provides an insight which indicates that the standard notion of KD nonclassicality [52] and the quantum advantage in postselected metrology may not have a direct connection.

TABLE I. KD distribution corresponding to the system state is tabulated. Under the optimality condition, the distribution is positive for all values of α. Here, $\{|\lambda\rangle, |\tilde{\lambda}\rangle, |-\lambda\rangle\}$ are the eigenvectors of system observable $\hat{A}$ and $\{|f_1\rangle, |f_2\rangle, |f_3\rangle\}$ is the complete set of postselection bases.

$\lim_{\phi\to\lambda\delta_\theta} \mathcal{K}^{\hat{\rho}_\theta}_{a_m,f_k}$	$\lvert f_1\rangle$	$\lvert f_2\rangle$	$\lvert f_3\rangle$
$\lvert\lambda\rangle$	0	$\frac{\cos^2\alpha}{2}$	$\frac{\sin^2\alpha}{2}$
$\lvert\tilde{\lambda}\rangle$	0	0	0
$\lvert-\lambda\rangle$	0	$\frac{\cos^2\alpha}{2}$	$\frac{\sin^2\alpha}{2}$

Rather the quantum advantage may have links with the nonvanishing quantum modification terms in KD distribution.

Thus, we have shown that the quantum advantage in postselected metrology does not appropriately reciprocate with the standard KD nonclassicality, which is often connected with negative or nonreal values in the distribution. Contrary to this, we have shown there exist states with positive KD distribution, which can lead the experimenter to have QFI beyond the HL in postselected metrology. Our results indicate that the quantum modification terms in the identity Eq. (36) effectively contribute to the efficiency in postselected metrology, which in consequence shows anomalous QFI.

V. DISCUSSION

We have studied postselected quantum metrology that has been shown to have better precision over other metrological protocols. It can even yield infinite precision, however, with very low probability. In a realistic situation, one has to consider both precision as well as probability. So far, an optimization protocol, which genuinely signifies the quantum advantages considering both of these quantities, has been missing.

We have shown that the accessible advantage is bounded irrespective of the quantum resources utilized in postselected metrology. Our results highlight the significance of weak values to study the quantum advantage. In general, the weak values and the geometric phases are interrelated [53,54]. We have provided a unique way to examine the optimality conditions by analyzing the geometric phases associated with the states under consideration. Using this scheme in postselected metrology, we have found an essential geometric argument in finding the optimal way to engineer the postselection.

It is evident that negative or nonreal elements in the Kirkwood-Dirac distribution are an essential signifier or indicator of nonclassicality. However, they do not represent a complete witness for nonclassicality. We prove that even positive Kirkwood-Dirac distributions can have operationally

TABLE II. Wigner formula corresponding to the system state.

$\lim_{\phi\to\lambda\delta_\theta} \mathcal{K}^{\hat{\rho}_\theta}_{a_m,f_k}$	$\lvert f_1\rangle$	$\lvert f_2\rangle$	$\lvert f_3\rangle$
$\lvert\lambda\rangle$	$\frac{1}{4}$	$\frac{\cos^2\alpha}{4}$	$\frac{\sin^2\alpha}{4}$
$\lvert\tilde{\lambda}\rangle$	0	0	0
$\lvert-\lambda\rangle$	$\frac{1}{4}$	$\frac{\cos^2\alpha}{4}$	$\frac{\sin^2\alpha}{4}$

relevant nonclassical features. However, the exact nonclassical feature of the Kirkwood-Dirac distribution that enables such advantage is still unknown. It can clearly be seen that the necessary and sufficient condition for which the efficiency of postselection metrology is zero is that all the intrinsic weak values must be independent of the postselection. The weak values are seen to be directly connected with conditional Kirkwood-Dirac quasiprobabilities, which are defined as Kirkwood-Dirac quasiprobabilities of a state subjected to a postselection. We conclude that the conditional KD distribution has to be independent of postselection (see Appendix G) in order to have zero quantum advantage. Classically, when conditional probabilities do not depend upon the postselection, it implies that the corresponding events are statistically independent. In this case, also, we can sense that there might be some statistical independence at the level of quasiprobabilities which prevents the experimenter from accessing the quantum advantage.

In summary, our work brings a better understanding of postselected metrology and provides a fundamental bound of quantum advantages irrespective of the quantum resources utilized. We provide a metrological protocol that can harness the maximum possible quantum advantage. Our results lay a foundation for further exploration of multiparameter postselected metrology and possible quantum advantages. Our study indicates that the negative or nonreal elements in the Kirkwood-Dirac distribution does not correspond to nonclassicality always. Thus, it is expected to initiate an effort for a complete characterization of nonclassicality on the theoretical ground.

ACKNOWLEDGMENT

M.N.B. gratefully acknowledges financial support from SERB-DST (Grant No. CRG/2019/002199), Government of India.

APPENDIX A: WEAK VALUE AMPLIFICATION

In standard weak value amplification (WVA) protocol, an experimenter prepares the meter and the system in some pure initial states $|\phi\rangle$ and $|\psi_i\rangle$, respectively. Then they are coupled weakly using the interaction Hamiltonian $\hat{H}_{\rm int} = \hbar g \hat{A} \otimes \hat{F}\,\delta(t - t_0)$, where $\hat{F}$ and $\hat{A}$ are observables corresponding to the meter and the system, respectively, and g is a small parameter signifying the coupling strength between the system and the meter. The function $\delta(t - t_0)$ indicates that the interaction between the system and the meter is impulsive. Finally, the postselection of the system is done onto a pure final state $|\psi_f\rangle$, discarding all the other events where the postselection fails. This procedure effectively prepares an updated meter state that includes the effect of the system $|\phi'\rangle = \hat{M}|\phi\rangle/||\hat{M}|\phi\rangle||$, which is mentioned in terms of the Kraus operator $\hat{M} = \langle\psi_f|\exp(-ig\hat{A}\otimes\hat{F})|\psi_i\rangle$. Averaging a meter observable $\hat{R}$ using this updated meter state yields $\langle\hat{R}\rangle_{|\phi'\rangle} = \langle\phi|\hat{M}^\dagger\hat{R}\hat{M}|\phi\rangle/\langle\phi|\hat{M}^\dagger\hat{M}|\phi\rangle$. The observable average is well approximated up to first order in g [55,56]:

$$\langle\hat{R}\rangle_{|\phi'\rangle} \approx 2g[\mathrm{Re}A_w\,\mathrm{Im}\alpha + \mathrm{Im}A_w\,\mathrm{Re}\alpha], \tag{A1}$$

where $\alpha = \langle\hat{R}\hat{F}\rangle_{|\phi\rangle}$ is the correlation parameter that can be fixed by the choice of meter observables and the initial meter state $|\phi\rangle$, and $A_w = \langle\psi_f|\hat{A}|\psi_i\rangle/\langle\psi_f|\psi_i\rangle$ is a complex *weak value* controlled by the system. This relation shows how a large weak value can enhance the sensitivity of the meter even with small g. In the case of estimating g, the weak value has notable ability to amplify its effect in the meter.

APPENDIX B: GENERALIZED PROTOCOL FOR OPTIMIZATION

All the optimization discussed so far can be generalized in terms of efficiency. At the point of optimization, the weak value and the postselection probability turn out to be

$$p_s = \frac{\langle\psi_i|\hat{A}|\psi_i\rangle^2}{\langle\psi_i|\hat{A}^2|\psi_i\rangle},\quad A_w = \frac{\langle\psi_i|\hat{A}^2|\psi_i\rangle}{\langle\psi_i|\hat{A}|\psi_i\rangle}. \tag{B1}$$

It is interesting to note that only the real part of the weak value will contribute when the point of saturation is reached. At this point, anomalously large weak values can be obtained when the observable average of $\hat{A}$ tends to zero. It has been shown that for a fixed weak value, the probability of successful postselection can be maximized by taking the final state parallel to $(\hat{A} - A_w)|\psi_i\rangle$ [57]. Alternatively, one can also conduct optimization to maximize the weak value for a fixed postselection probability. To accomplish such enhancements, one needs to use different optimization protocols. If we optimize the efficiency $\eta(A_w, p_s)$, it is still possible to attain the optimized values of the aforementioned protocols as special cases. Here, we demonstrate this conjunction of two different protocols with an example.

We consider coupling n entangled systems to the meter simultaneously where the system qubits are prepared in an entangled state [58],

$$|\psi_i\rangle = \frac{1}{\sqrt{2}}\left(|\lambda_x\rangle^{\otimes n} + |\lambda_y\rangle^{\otimes n}\right). \tag{B2}$$

Here, λ_x and λ_y are two arbitrary eigenvalues of an operator $\hat{a}$ whose eigenvectors span a subsystem. The rest of the subsystems are different copies of the same. From this preparation we construct the overall system observable as follows,

$$\hat{A} = \sum_{k=1}^{n} \mathbf{I}^{\otimes k-1} \otimes \hat{a} \otimes \mathbf{I}^{\otimes n-k}, \tag{B3}$$

where the operators in the tensor product act sequentially with different subsystems. The weak value, following Eq. (B1), is

$$A_w = \frac{n\left(\lambda_x^2 + \lambda_y^2\right)}{(\lambda_x + \lambda_y)}. \tag{B4}$$

Under the condition for a fixed weak value, we simply obtain

$$\lambda_y = \frac{A_w \pm \sqrt{A_w^2 - 4n^2\lambda_x^2 + 4nA_w\lambda_x}}{2n}. \tag{B5}$$

For large weak values the above expression can be approximated as $\lambda_y \approx \{(A_w - n\lambda_x)/2n, -\lambda_x\}$. The first root of λ_y corresponds to the case of nonanomalous weak value amplification. Under this choice, the eigenvalues of the operator are of the order of the weak value. However, the second choice, i.e., $\lambda_y \approx -\lambda_x$, corresponds to anomalous weak value

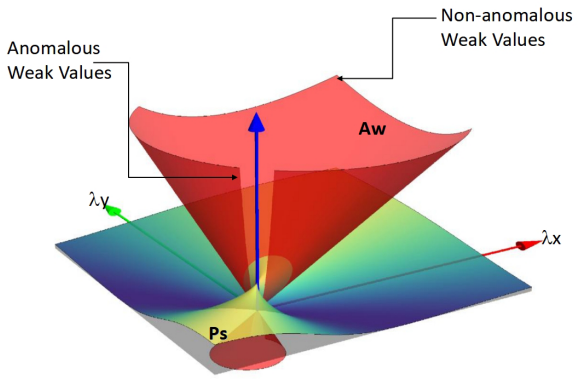

FIG. 3. Plot for optimal weak values and postselection probability for different initial state preparations characterized by λ_x and λ_y. The red surface corresponds to the optimal weak values A_w. The steep regions in this surface correspond to anomalous weak values, where they change rapidly upon small changes in the eigenvalues λ_x and λ_y due to anomalous amplification. In the nonanomalous region, the gradient is relatively low since the weak values scale with the eigenvalues here. The surface below stands for the corresponding success probability p_s. The blue regions stand for low success probability, and the yellow regions stand for high success probability.

amplification where the observable average with respect to the initial state approaches zero. The postselection probability for the optimal weak value turns out to be

$$p_s = n^2\lambda_x^2/A_w^2. \tag{B6}$$

Thus, the probability scales quadratically with n and shows improvement in estimation while comparing with the classical cases. Now, we explore the exact opposite case where the postselection probability is held fixed with a varying A_w. Starting with the same initial entangled state and using Eq. (B1), we arrive directly at

$$\lambda_y = \frac{2\lambda_x \pm \sqrt{4\lambda_x^2 - 4(2p_s-1)^2\lambda_x^2}}{4p_s - 2}. \tag{B7}$$

For very low p_s, this expression can be approximated as $\lambda_y \approx -\lambda_x$. Interestingly, we have traced the same approximated condition where anomalously large weak values appear (see Fig. 3). This yields the optimal weak value for a given probability of postselection,

$$|A_w| = n\lambda_x/\sqrt{p_s}, \tag{B8}$$

and it scales linearly with n. Clearly, the optimal values of A_w and p_s are appearing from a unified optimization process. In experiments, this unified optimization scheme would access large weak values at the cost of low probability and vice versa if the situation arises. Besides, this scheme also incorporates the cases where large eigenvalues of the system observable mostly do the amplification. This situation is referred to as the nonanomalous regime. In this regime, the weak values appear at the level of average values of the observable without any anomalous amplification. This would help us visualize the transition from weak values to average values, which is caused by the choice of preselection and postselection of the system in this scenario.

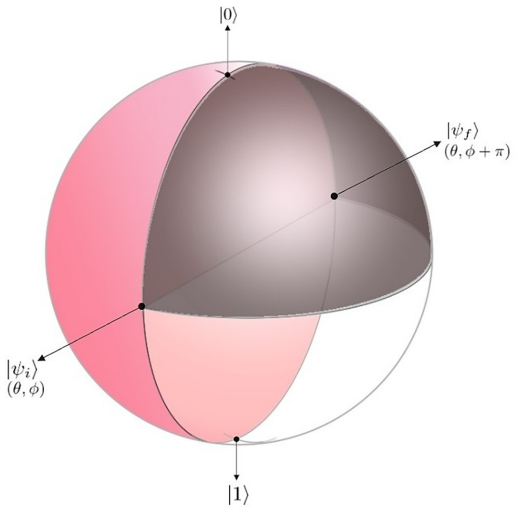

FIG. 4. Geometric phases associated with weak value amplification in the 2D Bloch sphere. The gray region is a lune in the 2D Bloch sphere with the dihedral angle θ; so it has area 2θ. This region is subtended by the state transition $|\psi_i\rangle \to |0\rangle \to |\psi_f\rangle \to |\psi_i\rangle$ which creates a geometric phase factor $\exp(-i\theta)$. The pink region is a half sphere with area 2π. A state transition $|\psi_i\rangle \xrightarrow{\text{via } |0\rangle} |1\rangle \xrightarrow{\text{via } |0\rangle} |\psi_f\rangle \to |\psi_i\rangle$ subtends the gray and the pink region together in Bloch sphere creating a geometric phase factor $-\exp(-i\theta)$. These two state transitions generate the optimal weak value amplification when a spin-1/2 system is prepared in $|\psi_i\rangle$ with latitude and longitude (θ, ϕ) in the 2D Bloch sphere and the system observable is σ_z.

APPENDIX C: EXAMPLE DEMONSTRATING THE GEOMETRIC OPTIMIZATION OF WVA

We take a spin-1/2 system as our reference for this example since the one qubit Bloch sphere is the simplest state-space geometry to visualize. The $\hat{\sigma}_z$ operator is taken as the system observable. Now we prepare the system in a state $|\psi_i\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$, which is represented as a point with latitude and longitude (θ, ϕ) in the Bloch sphere. The north and south poles refer to the $|0\rangle$ and $|1\rangle$ states, respectively.

Now, the optimal efficiency in weak value is attained when the system is postselected in $|\psi_f\rangle = \cos(\theta/2)|0\rangle - e^{i\phi}\sin(\theta/2)|1\rangle$, which specifically refers to the point in the Bloch sphere with coordinates $(\theta, \phi + \pi)$. Since in the optimal limit the geometric phases $\Phi_g^{a_k}$ depend on the eigenvalues of the observable via sign functions, we can expect the following in the optimal limit:

$$\exp\left[i\Phi_g^1(|\psi_i\rangle_{\theta,\phi}, |0\rangle, |\psi_f\rangle_{\theta,\phi+\pi})\right] = \chi, \tag{C1}$$

$$\exp\left[i\Phi_g^{-1}(|\psi_i\rangle_{\theta,\phi}, |1\rangle, |\psi_f\rangle_{\theta,\phi+\pi})\right] = -\chi, \tag{C2}$$

where χ is some constant independent of the eigenvalues of the observable. The gray region in Fig. 4 is subtended by the transition path $|\psi_i\rangle \to |0\rangle \to |\psi_f\rangle \to |\psi_i\rangle$. All the transitions are executed through great circles in the 2D Bloch sphere. The area of this region is 2θ, so this region will create a geometric phase factor $\exp(-i\theta)$. This geometric phase corresponds to the first state transition in Eq. (C1). We take the second transition $|\psi_i\rangle \xrightarrow{\text{via } |0\rangle} |1\rangle \xrightarrow{\text{via } |0\rangle} |\psi_f\rangle \to |\psi_i\rangle$. This transition

subtends the black and the pink region together. The pink region is a half sphere, so it has area 2π. Thus, the total area covered by the second transition is $2\pi + 2\theta$. This creates a total geometric phase factor $\exp(-i(2\pi + 2\theta)/2)$ which equals $-\exp(-i\theta)$. Thus, taking $\chi = \exp(-i\theta)$ reconciles this geometric understanding of weak value optimization with the standard weak value optimization procedure.

APPENDIX D: RELATION BETWEEN POSTSELECTED FISHER INFORMATION AND WEAK VALUES

The postselected quantum Fisher information is given by

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big) = 4\langle\dot\Psi_\theta^{\rm ps}|\dot\Psi_\theta^{\rm ps}\rangle\frac{1}{p_\theta^{\rm ps}} - 4\big|\langle\dot\Psi_\theta^{\rm ps}|\Psi_\theta^{\rm ps}\rangle\big|^2\frac{1}{\big(p_\theta^{\rm ps}\big)^2}, \quad \text{(D1)}$$

where the unnormalized postselected quantum state is $|\Psi_\theta^{\rm ps}\rangle = \hat F\hat U(\theta)|\psi_0\rangle$, where $|\psi_0\rangle\langle\psi_0| \equiv \hat\rho_0$. $p_\theta^{\rm ps} = {\rm Tr}(\hat F\hat\rho_\theta)$ is the probability of postselection. In this Appendix, we show that how the postselected Fisher information and weak values are related. Before doing this, we quickly recap the way of re-expressing Eq. (D1) in terms of the operator form. The first term of Eq. (D1) is

$$\frac{4}{p_\theta^{\rm ps}}\langle\dot\Psi_\theta^{\rm ps}|\dot\Psi_\theta^{\rm ps}\rangle = \frac{4}{p_\theta^{\rm ps}}{\rm Tr}[\hat F\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}^\dagger(\theta)\hat F^\dagger] = \frac{4}{p_\theta^{\rm ps}}{\rm Tr}(\hat F\hat A\hat\rho_\theta\hat A). \quad \text{(D2)}$$

We can express $\hat A$ and $\hat F$ in their corresponding eigendecompositions to rewrite Eq. (D2) in terms of the doubly extended Kirkwood-Dirac quasiprobability distribution. Similarly, the second term of Eq. (D1) is

$$\frac{4}{\big(p_\theta^{\rm ps}\big)^2}\big|\langle\dot\Psi_\theta^{\rm ps}|\Psi_\theta^{\rm ps}\rangle\big|^2 = \frac{4}{\big(p_\theta^{\rm ps}\big)^2}|{\rm Tr}(\hat F\hat\rho_\theta\hat A)|^2. \quad \text{(D3)}$$

Combining the expressions above, we obtain

$$\mathcal{F}_Q\big(\theta\big|\psi_\theta^{\rm ps}\big) = \frac{4}{p_\theta^{\rm ps}}{\rm Tr}[\hat F\hat A\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger\hat A] - \frac{4}{\big(p_\theta^{\rm ps}\big)^2}|{\rm Tr}[\hat F\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger\hat A]|^2. \quad \text{(D4)}$$

Rearranging Eq. (8) in the main text along with the replacement of total probability $p_\theta^{\rm ps}$ as a sum of individual probabilities when the postselction is done with a rank-2 projector, we get

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4[\big(p_{\theta1}^{\rm ps} + p_{\theta2}^{\rm ps}\big)(|\langle\psi_\theta|\hat A|f_1\rangle|^2 + |\langle\psi_\theta|\hat A|f_2\rangle|^2) - |\langle\psi_\theta|\hat A|f_1\rangle\langle f_1|\psi_\theta\rangle + \langle\psi_\theta|\hat A|f_2\rangle\langle f_2|\psi_\theta\rangle|^2]. \quad \text{(D5)}$$

Simplifying Eq. (D5), we find

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4\big[p_{\theta1}^{\rm ps}|\langle\psi_\theta|\hat A|f_1\rangle|^2 + p_{\theta2}^{\rm ps}|\langle\psi_\theta|\hat A|f_2\rangle|^2 + p_{\theta2}^{\rm ps}|\langle\psi_\theta|\hat A|f_1\rangle|^2 + p_{\theta1}^{\rm ps}|\langle\psi_\theta|\hat A|f_2\rangle|^2 - p_{\theta1}^{\rm ps}|\langle\psi_\theta|\hat A|f_1\rangle|^2 - p_{\theta2}^{\rm ps}|\langle\psi_\theta|\hat A|f_2\rangle|^2 - \langle\psi_\theta|\hat A|f_1\rangle\langle f_1|\psi_\theta\rangle\langle\psi_\theta|f_2\rangle\langle f_2|\hat A|\psi_\theta\rangle - \langle\psi_\theta|\hat A|f_2\rangle\langle f_2|\psi_\theta\rangle\langle\psi_\theta|f_1\rangle\langle f_1|\hat A|\psi_\theta\rangle\big]. \quad \text{(D6)}$$

This expression can be written in a more concise way:

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4\big[p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}\big|A_w^1\big|^2 + p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}\big|A_w^2\big|^2 - p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}\ A_w^2A_w^{1*} - p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}A_w^1A_w^{2*}\big], \quad \text{(D7)}$$

where $A_w^k = \frac{\langle\psi_\theta|\hat A|f_k\rangle}{\langle\psi_\theta|f_k\rangle}$. Further simplification leads to

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}\big|A_w^1 - A_w^2\big|^2. \quad \text{(D8)}$$

Repeating the same calculation for a rank-3 projector, we obtain

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4\big[p_{\theta1}^{\rm ps}p_{\theta2}^{\rm ps}\big|A_w^1 - A_w^2\big|^2 + p_{\theta2}^{\rm ps}p_{\theta3}^{\rm ps}\big|A_w^2 - A_w^3\big|^2 + p_{\theta3}^{\rm ps}p_{\theta1}^{\rm ps}\big|A_w^3 - A_w^1\big|^2\big]. \quad \text{(D9)}$$

From these two expressions we can easily extrapolate a generalized formula for a rank-n projector as follows:

$$\mathcal{F}_Q\big(\theta|\psi_\theta^{\rm ps}\big)\big(p_\theta^{\rm ps}\big)^2 = 4\sum_{m<n}p_{\theta m}^{\rm ps}p_{\theta n}^{\rm ps}\big|A_w^m - A_w^n\big|^2. \quad \text{(D10)}$$

APPENDIX E: EFFICIENCY OF THE POSTSELECTED METROLOGY

We begin by writing the the evolved pure state in the following form,

$$\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger = |\psi_\theta\rangle\langle\psi_\theta|. \quad \text{(E1)}$$

Rewriting the first part of the right-hand side in Eq. (8) in the main text followed by a decomposition for the postselection projectors yield, we get

$${\rm Tr}[\hat F\hat A\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger\hat A] = \sum_{f_k\in\mathcal{F}^{\rm ps}}|\langle\psi_\theta|\hat A|f_k\rangle|^2. \quad \text{(E2)}$$

Using weak value optimization and rewriting the previous equation, we get

$${\rm Tr}[\hat F\hat A\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger\hat A] = \sum_{f_k\in\mathcal{F}^{\rm ps}}p_{\theta k}^{\rm ps}\big|A_w^k\big|^2. \quad \text{(E3)}$$

This expression establishes a connection between the efficiency of weak value amplification and the efficiency of postselected quantum metrology. Using a similar method we can derive the following regarding the second term in the the right-hand side of Eq. (40) in the main text:

$${\rm Tr}[\hat F\dot{\hat U}(\theta)\hat\rho_0\dot{\hat U}(\theta)^\dagger\hat A] = \sum_{f_k\in\mathcal{F}^{\rm ps}}\langle\psi_\theta|\hat A|f_k\rangle\langle f_k|\psi_\theta\rangle. \quad \text{(E4)}$$

Rewriting the same equation after a spectral decomposition of $\hat{A}$, we obtain

$$\begin{aligned}\mathrm{Tr}[\hat{F}\hat{U}(\theta)\hat{\rho}_0\hat{U}(\theta)^\dagger\hat{A}] &= \sum_m \sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m\langle\psi_\theta|a_m\rangle\langle a_m|f_k\rangle\langle f_k|\psi_\theta\rangle \\ &= \sum_m \sum_{f_k\in\mathcal{F}^{\mathrm{ps}}} a_m q^{(a_m,f_k)}_{|\psi_\theta\rangle}. \end{aligned} \tag{E5}$$

Putting these expressions altogether, we can easily obtain Eq. (21) in the main text.

APPENDIX F: DERIVATION OF THE RELATIONSHIP BETWEEN EFFICIENCY OF WVA AND GEOMETRIC PHASE

We start with the expression for the geometric phase,

$$\begin{aligned}\Phi_{\mathrm{g}}^{a_k}(|\psi_i\rangle,|a_k\rangle,|\psi_f\rangle) &= \arg(\langle\psi_f|a\rangle\langle a|\psi_i\rangle\langle\psi_i|\psi_f\rangle) \\ &= \arg(\langle\psi_f|a\rangle\langle a|\psi_i\rangle) + \arg(\langle\psi_i|\psi_f\rangle). \end{aligned} \tag{F1}$$

Rearranging Eq. (F1), we obtain

$$\arg(\langle\psi_f|a\rangle\langle a|\psi_i\rangle) = \Phi_{\mathrm{g}}^{a_k}(|\psi_i\rangle,|a_k\rangle,|\psi_f\rangle) - \arg(\langle\psi_i|\psi_f\rangle). \tag{F2}$$

Rewriting the efficiency of WVA in terms of the geometric phase, we get

$$\begin{aligned}\eta(p_s,A_w) &= |\langle\psi_f|\hat{A}|\psi_i\rangle|^2 \\ &= \left|\sum_k a_k\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle\right|^2 \\ &= \left|\sum_k a_k\big|\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle\big|\right. \\ &\quad \left.\times\exp[i\arg(\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle)]\right|^2. \end{aligned} \tag{F3}$$

Replacing Eq. (F2) in Eq. (F3), we get

$$\begin{aligned}\eta(p_s,A_w) = \Bigg|\sum_k a_k|\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle| \\ \times\exp\Big\{i\left[\Phi_{\mathrm{g}}^{a_k}(|\psi_i\rangle,|a_k\rangle,|\psi_f\rangle)-\arg(\langle\psi_i|\psi_f\rangle)\right]\Big\}\Bigg|^2. \end{aligned} \tag{F4}$$

Since $\langle\psi_i|\psi_f\rangle$ is not dependent on the eigenvalues of the system observable, it creates a global phase and does not contribute to the efficiency. Thus, we obtain

$$\eta(A_w,p_s) = \Big|\sum_k a_k\big|\langle\psi_f|a_k\rangle\langle a_k|\psi_i\rangle\big|\exp\big(i\,\Phi_g^{a_k}\big)\Big|^2. \tag{F5}$$

APPENDIX G: CONDITIONAL KD DISTRIBUTION AND EFFICIENCY IN POSTSELECTED METROLOGY

We start with Eq. (10) and see that the necessary and sufficient condition to have $\xi(p_\theta^{\mathrm{ps}};\mathcal{F}_Q)=0$ is that all the intrinsic weak values A_w^k are independent of postselection f_k for all $f_k\in\mathcal{F}^{\mathrm{ps}}$. A natural question would arise here: how is this condition being manifested at the level of Kirkwood-Dirac quasiprobabilities? To answer this, we refer readers to see the discussion on the conditional Kirkwood-Dirac distribution [59,60]: a distribution of a state ($|\psi\rangle$) that is conditioned on postselection ($|f\rangle\langle f|$), $\mathcal{KC}^{(a_m,b_n|f)}_{|\psi\rangle} := \frac{\langle\psi|a_m\rangle\langle a_m|b_n\rangle\langle b_n|\psi\rangle}{|\langle\psi|f\rangle|^2}$, where $\{|a_m\rangle\}$ and $\{|b_n\rangle\}$ are two sets of complete orthonormal bases referring to the points in the distribution. Note that, when postselection $|f\rangle$ belongs to the orthonormal set $\{|b_n\rangle\}$, the conditional KD distribution equals a weak value of a projector $|a_m\rangle\langle a_m|$, $\mathcal{KC}^{(a_m,f|f)}_{|\psi\rangle} = \frac{\langle\psi|a_m\rangle\langle a_m|f\rangle}{\langle\psi|f\rangle}$. We now prove that all the weak values A_w^k are independent irrespective of postselection, which straightforwardly implies that $\mathcal{KC}^{(a,f_k|f_k)}_{|\psi_\theta\rangle}$ must be independent of the postselection f_k. $A_w^k = A_w^l, \forall f_k, f_l\in\mathcal{F}^{\mathrm{ps}} \iff \sum_a a(\mathcal{KC}^{(a,f_k|f_k)}_{|\psi_\theta\rangle} - \mathcal{KC}^{(a,f_l|f_l)}_{|\psi_\theta\rangle}) = 0$. It can be seen that any shift [39] in the eigenvalues of $\hat{A}$ does not contribute to the Fisher information for the quasiprobabilities which are potent to harness nonclassical advantage. Thus, we can assume the eigenvalues a are all positive, which implies that $\mathcal{KC}^{(a,f_k|f_k)}_{|\psi_\theta\rangle} = \mathcal{KC}^{(a,f_l|f_l)}_{|\psi_\theta\rangle}, \forall f_k, f_l\in\mathcal{F}^{\mathrm{ps}}$.

[1] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*, Vol. 9 (Princeton University, Princeton, NJ, 2016).

[2] C. R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, in *Breakthroughs in Statistics* (Springer, Berlin, 1992), pp. 235–247.

[3] V. Giovannetti, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306**, 1330 (2004).

[4] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[5] M. G. Paris, Quantum estimation for quantum technology, Int. J. Quantum Inform. **07**, 125 (2009).

[6] B. R. Frieden, Physics from Fisher information: A unification, Am. J. Phys. **68**, 1064 (2000).

[7] D. Petz and C. Ghinea, Introduction to quantum Fisher information, in *Quantum Probability and Related Topics* (World Scientific, Singapore, 2011), pp. 261–281.

[8] L. Maccone, Intuitive reason for the usefulness of entanglement in quantum metrology, Phys. Rev. A **88**, 042109 (2013).

[9] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, Phys. Rev. Lett. **96**, 010401 (2006).

[10] M. Zwierz, C. A. Pérez-Delgado, and P. Kok, Ultimate limits to quantum metrology and the meaning of the Heisenberg limit, Phys. Rev. A **85**, 042112 (2012).

[11] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D **23**, 1693 (1981).

[12] J. P. Dowling, Quantum optical metrology—the lowdown on high-N00N states, Contemp. Phys. **49**, 125 (2008).

[13] H. Vahlbruch, S. Chelkowski, B. Hage, A. Franzen, K. Danzmann, and R. Schnabel, Coherent Control of Vacuum Squeezing in the Gravitational-Wave Detection Band, Phys. Rev. Lett. **97**, 011101 (2006).

[14] L. Barsotti, J. Harms, and R. Schnabel, Squeezed vacuum states of light for gravitational wave detectors, Rep. Prog. Phys. **82**, 016905 (2019).

[15] K. Goda, O. Miyakawa, E. E. Mikhailov, S. Saraf, R. Adhikari, K. McKenzie, R. Ward, S. Vass, A. J. Weinstein, and N. Mavalvala, A quantum-enhanced prototype gravitational-wave detector, Nat. Phys. **4**, 472 (2008).

[16] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, Noisy Metrology beyond the Standard Quantum Limit, Phys. Rev. Lett. **111**, 120401 (2013).

[17] M. B. Plenio and S. F. Huelga, Sensing in the presence of an observed environment, Phys. Rev. A **93**, 032123 (2016).

[18] F. Albarelli, M. A. Rossi, M. G. Paris, and M. G. Genoni, Ultimate limits for quantum magnetometry via time-continuous measurements, New J. Phys. **19**, 123011 (2017).

[19] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum Metrology in Non-Markovian Environments, Phys. Rev. Lett. **109**, 233601 (2012).

[20] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański, Ultimate Precision Limits for Noisy Frequency Estimation, Phys. Rev. Lett. **116**, 120801 (2016).

[21] N. Mauranyapin, L. Madsen, M. Taylor, M. Waleed, and W. Bowen, Evanescent single-molecule biosensing with quantum-limited precision, Nat. Photonics **11**, 477 (2017).

[22] M. A. Taylor and W. P. Bowen, Quantum metrology and its application in biology, Phys. Rep. **615**, 1 (2016).

[23] LIGO Scientific Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, C. Adams, R. Adhikari, C. Affeldt, B. Allen, G. Allen *et al.*, A gravitational wave observatory operating beyond the quantum shot-noise limit, Nat. Phys. **7**, 962 (2011).

[24] J. Borregaard and A. S. Sørensen, Near-Heisenberg-Limited Atomic Clocks in the Presence of Decoherence, Phys. Rev. Lett. **111**, 090801 (2013).

[25] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt, Optical atomic clocks, Rev. Mod. Phys. **87**, 637 (2015).

[26] H. Katori, Optical lattice clocks and quantum metrology, Nat. Photonics **5**, 203 (2011).

[27] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, Optics and interferometry with atoms and molecules, Rev. Mod. Phys. **81**, 1051 (2009).

[28] Y. Che, J. Liu, X.-M. Lu, and X. Wang, Multiqubit matter-wave interferometry under decoherence and the Heisenberg scaling recovery, Phys. Rev. A **99**, 033807 (2019).

[29] D. Tsarev, S. Arakelian, Y.-L. Chuang, R.-K. Lee, and A. Alodjants, Quantum metrology beyond Heisenberg limit with entangled matter wave solitons, Opt. Express **26**, 19583 (2018).

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

[31] J. Zhang and M. Sarovar, Quantum Hamiltonian Identification from Measurement Time Traces, Phys. Rev. Lett. **113**, 080401 (2014).

[32] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, Estimation of many-body quantum Hamiltonians via compressive sensing, Phys. Rev. A **84**, 012107 (2011).

[33] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Hamiltonian Learning and Certification Using Quantum Resources, Phys. Rev. Lett. **112**, 190501 (2014).

[34] T. Gefen, A. Rotem, and A. Retzker, Overcoming resolution limits with quantum sensing, Nat. Commun. **10**, 4992 (2019).

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

[36] M. Napolitano, M. Koschorreck, B. Dubost, N. Behbood, R. Sewell, and M. W. Mitchell, Interaction-based quantum metrology showing scaling beyond the Heisenberg limit, Nature (London) **471**, 486 (2011).

[37] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Generalized Limits for Single-Parameter Quantum Estimation, Phys. Rev. Lett. **98**, 090401 (2007).

[38] S. Choi and B. Sundaram, Bose-Einstein condensate as a nonlinear Ramsey interferometer operating beyond the Heisenberg limit, Phys. Rev. A **77**, 053613 (2008).

[39] D. R. M. Arvidsson-Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Quantum advantage in postselected metrology, Nat. Commun. **11**, 3775 (2020).

[40] P. A. M. Dirac, On the analogy between classical and quantum mechanics, Rev. Mod. Phys. **17**, 195 (1945).

[41] J. G. Kirkwood, Quantum statistics of almost classical assemblies, Phys. Rev. **44**, 31 (1933).

[42] L. M. Johansen, Quantum theory of successive projective measurements, Phys. Rev. A **76**, 012119 (2007).

[43] R. W. Spekkens, Negativity and Contextuality Are Equivalent Notions of Nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[44] A. Barut, Distribution functions for noncommuting operators, Phys. Rev. **108**, 565 (1957).

[45] Y. Terletsky, On the classical limit of quantum mechanics, Zh. Eksp. Teor. Fiz. **7**, 1290 (1937).

[46] H. Margenau and R. N. Hill, Correlation between measurements in quantum theory, Prog. Theor. Phys. **26**, 722 (1961).

[47] S. Tanaka and N. Yamamoto, Information amplification via postselection: A parameter-estimation perspective, Phys. Rev. A **88**, 042116 (2013).

[48] J. Combes, C. Ferrie, Z. Jiang, and C. M. Caves, Quantum limits on postselected, probabilistic quantum metrology, Phys. Rev. A **89**, 052117 (2014).

[49] These weak values are not directly measured through standard weak measurement. Rather, we see their effects indirectly as they play a key role in determining the postselected QFI. We refer to them as the intrinsic weak values of the observable.

[50] N. Mukunda and R. Simon, Quantum kinematic approach to the geometric phase. I. General formalism, Ann. Phys. **228**, 205 (1993).

[51] G. B. Alves, B. M. Escher, R. L. de Matos Filho, N. Zagury, and L. Davidovich, Weak-value amplification as an optimal metrological protocol, Phys. Rev. A **91**, 062107 (2015).

[52] The standard notion of KD nonclassicality refers negative or nonreal entries in KD distribution.

[53] E. Sjöqvist, Geometric phase in weak measurements, Phys. Lett. A **359**, 187 (2006).

[54] S. Tamate, H. Kobayashi, T. Nakanishi, K. Sugiyama, and M. Kitano, Geometrical aspects of weak measurements and quantum erasers, New J. Phys. **11**, 093025 (2009).

[55] A. G. Kofman, S. Ashhab, and F. Nori, Nonperturbative theory of weak pre-and postselected measurements, Phys. Rep. **520**, 43 (2012).

[56] A. Di Lorenzo, Full counting statistics of weak-value measurement, Phys. Rev. A **85**, 032106 (2012).

[57] S. Pang, J. Dressel, and T. A. Brun, Entanglement-Assisted Weak Value Amplification, Phys. Rev. Lett. **113**, 030401 (2014).

[58] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).

[59] L. M. Johansen, Nonclassical properties of coherent states, Phys. Lett. A **329**, 184 (2004).

[60] N. Yunger Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, Phys. Rev. A **97**, 042105 (2018).