

# Quantum Advantage in Postselected Metrology

David R. M. Arvidsson-Shukur,<sup>1,2,3</sup> Nicole Yunger Halpern,<sup>4,5,3</sup> Hugo V. Lepage,<sup>6</sup> Aleksander A. Lasek,<sup>6</sup> Crispin H. W. Barnes,<sup>6</sup> and Seth Lloyd<sup>2,3</sup>

<sup>1</sup>*Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, United Kingdom*

<sup>2</sup>*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

<sup>3</sup>*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

<sup>4</sup>*ITAMP, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA*

<sup>5</sup>*Department of Physics, Harvard University, Cambridge, MA 02138, USA*

<sup>6</sup>*Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, United Kingdom*

(Dated: February 9, 2021)

We show that postselection offers a nonclassical advantage in metrology. In every parameter-estimation experiment, the final measurement or the postprocessing incurs some cost. Postselection can improve the rate of Fisher information (the average information learned about an unknown parameter from an experimental trial) to cost. This improvement, we show, stems from the negativity of a quasiprobability distribution, a quantum extension of a probability distribution. In a classical theory, in which all observables commute, our quasiprobability distribution can be expressed as real and nonnegative. In a quantum-mechanically noncommuting theory, nonclassicality manifests in negative or nonreal quasiprobabilities. The distribution's nonclassically negative values enable postselected experiments to outperform even postselection-free experiments whose input states and final measurements are optimized: Postselected quantum experiments can yield anomalously large information-cost rates. We prove that this advantage is genuinely nonclassical: no classically commuting theory can describe any quantum experiment that delivers an anomalously large Fisher information. Finally, we outline a preparation-and-postselection procedure that can yield an *arbitrarily* large Fisher information. Our results establish the nonclassicality of a metrological advantage, leveraging our quasiprobability distribution as a mathematical tool.

## INTRODUCTION

Our ability to deliver new quantum-mechanical improvements to technologies relies on a better understanding of the foundation of quantum theory: When is a phenomenon truly nonclassical? We take noncommutation as our notion of nonclassicality and we quantify this nonclassicality with negativity: Quantum states can be represented by quasiprobability distributions, extensions of classical probability distributions. Whereas probabilities are real and nonnegative, quasiprobabilities can assume negative and nonreal values. Quasiprobabilities' negativity stems from the impossibility of representing quantum states with joint probability distributions [1–3]. The distribution we use, an extension of the Kirkwood-Dirac distribution [4–6], signals nonclassical noncommutation through the presence of negative or nonreal quasiprobabilities.

One field advanced by quantum mechanics is metrology, which concerns the statistical estimation of unknown physical parameters. Quantum metrology relies on quantum phenomena to improve estimations beyond classical bounds [7]. A famous example exploits entanglement [8–10]. Consider using  $N$  separable and distinguishable probe states to evaluate identical systems in parallel. The best estimator's error will scale as  $N^{-1/2}$ . If the probes are entangled, the error scaling improves to  $N^{-1}$  [11]. As Bell's theorem rules out classical (local realist) explanations of entanglement, the improvement is genuinely quantum.

A central quantity in parameter estimation is the

Fisher information,  $\mathcal{I}(\theta)$ . The Fisher information quantifies the average information learned about an unknown parameter  $\theta$  from an experiment [12–14].  $\mathcal{I}(\theta)$  lower-bounds the variance of an unbiased estimator  $\theta_e$  via the Cramér-Rao inequality:  $\text{Var}(\theta_e) \geq 1/\mathcal{I}(\theta)$  [15, 16]. A common metrological task concerns optimally estimating a parameter that characterizes a physical process. The experimental input and the final measurement are optimized to maximize the Fisher information and to minimize the estimator's error.

Classical parameter estimation can benefit from postselecting the output data before postprocessing. Postselection can raise the Fisher information per final measurement or postprocessing event (Fig. 1). Postselection can also raise the rate of information per final measurement in a quantum setting. But classical postselection is intuitive, whereas an intense discussion surrounds postselected quantum experiments [17–30]. The ontological nature of postselected quantum states, and the extent to which they exhibit nonclassical behavior, is subject to an ongoing debate. Particular interest has been aimed at pre- and postselected averages of observables. These *weak values* can lie outside an observable's eigenspectrum when measured via a weak coupling to a pointer particle [17, 31]. Such values offer metrological advantages in estimations of weak-coupling strengths [19, 24, 32–37].

In this article, we go beyond this restrictive setting and ask, can postselection provide a nonclassical advantage in *general* quantum parameter-estimation experiments? We conclude that it can. We study metrology experiments for estimating an unknown transfor-

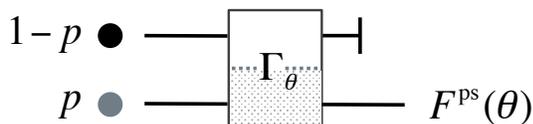


FIG. 1. **Classical experiment with postselection.** A nonoptimal input device initializes a particle in one of two states, with probabilities  $p$  and  $1 - p$ , respectively. The particle undergoes a transformation  $\Gamma_\theta$  set by an unknown parameter  $\theta$ . Only the part of the transformation that acts on particles in the lower path depends on  $\theta$ . If the final measurement is expensive, the particles in the upper path should be discarded: they possess no information about  $\theta$ .

mation parameter whose final measurement or postprocessing incurs an experimental cost [38, 39]. Postselection allows the experiment to incur that cost only when the postselected measurement's result reveals that the final measurement's Fisher information will be sufficiently large. We express the Fisher information in terms of a quasiprobability distribution. Quantum negativity in this distribution enables postselection to increase the Fisher information above the values available from standard input-and-measurement-optimized experiments. Such an anomalous Fisher information can improve the rate of information gain to experimental cost, offering a genuine quantum advantage in metrology. We show that, within a commuting theory, a theory in which observables commute classically, postselection can improve information-cost rates no more than a strategy that uses an optimal input and final measurement can. We thus conclude that experiments that generate anomalous Fisher-information values require noncommutativity.

## RESULTS

### Postselected quantum Fisher information

As aforementioned, postselection can raise the Fisher information per final measurement. Figure 1 outlines a classical experiment with such an information enhancement. Below, we show how postselection affects the Fisher information in a quantum setting.

Consider an experiment with outcomes  $i$  and associated probabilities  $p_i(\theta)$ , which depend on some unknown parameter  $\theta$ . The Fisher information about  $\theta$  is [14]

$$\mathcal{I}(\theta) = \sum_i p_i(\theta) [\partial_\theta \ln(p_i(\theta))]^2 = \sum_i \frac{1}{p_i(\theta)} [\partial_\theta p_i(\theta)]^2. \quad (1)$$

Repeating the experiment  $N \gg 1$  times provides, on average, an amount  $N\mathcal{I}(\theta)$  of information about  $\theta$ . The estimator's variance is bounded by  $\text{Var}(\theta_e) \geq 1/[N\mathcal{I}(\theta)]$ .

Below, we define and compare two types of metrological procedures. In both scenarios, we wish to estimate

an unknown parameter  $\theta$  that governs a physical transformation.

**Optimized prepare-measure experiment:** An input system undergoes the partially unknown transformation, after which the system is measured. Both the input system and the measurement are chosen to provide the largest possible Fisher information.

**Postselected prepare-measure experiment:** An input system undergoes, first, the partially unknown transformation and, second, a postselection measurement. Conditioned on the postselection's yielding the desired outcome, the system undergoes an information-optimized final measurement.

In quantum parameter estimation, a quantum state is measured to reveal information about an unknown parameter encoded in the state. We now compare, in this quantum setting, the Fisher-information values generated from the two metrological procedures described above. Consider a quantum experiment that outputs a state  $\hat{\rho}_\theta = \hat{U}(\theta)\hat{\rho}_0\hat{U}^\dagger(\theta)$ , where  $\hat{\rho}_0$  is the input state and  $\hat{U}(\theta)$  represents a unitary evolution set by  $\theta$ . The quantum Fisher information is defined as the Fisher information maximized over all possible generalized measurements [7, 13, 40–42]:

$$\mathcal{I}_Q(\theta|\hat{\rho}_\theta) = \text{Tr}[\hat{\rho}_\theta \hat{\Lambda}_{\hat{\rho}_\theta}^2]. \quad (2)$$

$\hat{\Lambda}_{\hat{\rho}_\theta}$  is the symmetric logarithmic derivative, implicitly defined by  $\partial_\theta \hat{\rho}_\theta = \frac{1}{2}(\hat{\Lambda}_{\hat{\rho}_\theta} \hat{\rho}_\theta + \hat{\rho}_\theta \hat{\Lambda}_{\hat{\rho}_\theta})$  [12].

If  $\hat{\rho}_\theta$  is pure, such that  $\hat{\rho}_\theta = |\Psi_\theta\rangle\langle\Psi_\theta|$ , the quantum Fisher information can be written as [33, 43]

$$\mathcal{I}_Q(\theta|\hat{\rho}_\theta) = 4 \langle \dot{\Psi}_\theta | \dot{\Psi}_\theta \rangle - 4 | \langle \dot{\Psi}_\theta | \Psi_\theta \rangle |^2, \quad (3)$$

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

We assume that the evolution can be represented in accordance with Stone's theorem [44], by  $\hat{U}(\theta) \equiv e^{-i\hat{A}\theta}$ , where  $\hat{A}$  is a Hermitian operator. We assume that  $\hat{A}$  is not totally degenerate: If all the  $\hat{A}$  eigenvalues were identical,  $\hat{U}(\theta)$  would not imprint  $\theta$  onto the state in a relative phase. For a pure state, the quantum Fisher information equals  $\mathcal{I}_Q(\theta|\hat{\rho}_\theta) = 4\text{Var}(\hat{A})_{\hat{\rho}_0}$  [7]. Maximizing Eq. 1 over all measurements gives  $\mathcal{I}_Q(\theta|\hat{\rho}_\theta)$ . Similarly,  $\mathcal{I}_Q(\theta|\hat{\rho}_\theta)$  can be maximized over all input states. For a given unitary  $\hat{U}(\theta) = e^{-i\hat{A}\theta}$ , the maximum quantum Fisher information is

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

where  $\Delta a$  is the difference between the maximum and minimum eigenvalues of  $\hat{A}$  [7].<sup>1</sup> To summarize, in an optimized quantum prepare-measure experiment, the quantum Fisher information is  $(\Delta a)^2$ .

<sup>1</sup> The information-optimal input state is a pure state in an equal superposition of one eigenvector associated with the smallest eigenvalue and one associated with the largest.

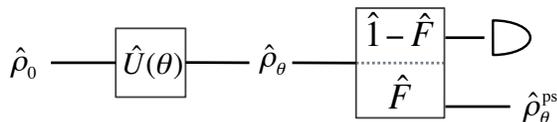


FIG. 2. **Preparation of postselected quantum state.** First, an input quantum state  $\hat{\rho}_0$  undergoes a unitary transformation  $\hat{U}(\theta) = e^{-i\theta\hat{A}}$ :  $\hat{\rho}_0 \rightarrow \hat{\rho}_\theta$ . Second, the quantum state is subject to a projective postselective measurement  $\{\hat{F}, \hat{1} - \hat{F}\}$ . The postselection is such that if the outcome related to the operator  $\hat{F}$  happens, then the quantum state is not destroyed. The experiment outputs renormalized states  $\hat{\rho}_\theta^{\text{ps}} = \hat{F}\hat{\rho}_\theta\hat{F}/\text{Tr}(\hat{F}\hat{\rho}_\theta)$ .

We now find an expression for the quantum Fisher information in a postselected prepare-measure experiment. A projective postselection occurs after  $\hat{U}(\theta)$  but before the final measurement. Figure 2 shows such a quantum circuit. The renormalized quantum state that passes the postselection is  $|\Psi_\theta^{\text{ps}}\rangle \equiv |\psi_\theta^{\text{ps}}\rangle/\sqrt{p_\theta^{\text{ps}}}$ , where we have defined an unnormalized state  $|\psi_\theta^{\text{ps}}\rangle \equiv \hat{F}|\Psi_\theta\rangle$  and the postselection probability  $p_\theta^{\text{ps}} \equiv \text{Tr}(\hat{F}\hat{\rho}_\theta)$ . As before,  $\hat{\rho}_\theta = \hat{U}(\theta)\hat{\rho}_0\hat{U}^\dagger(\theta)$ .  $\hat{F} = \sum_{f \in \mathcal{F}^{\text{ps}}} |f\rangle\langle f|$  is the postselecting projection operator, and  $\mathcal{F}^{\text{ps}}$  is a set of orthonormal basis states allowed by the postselection. Finally, the postselected state undergoes an information-optimal measurement.

When  $|\Psi_\theta^{\text{ps}}\rangle \equiv |\psi_\theta^{\text{ps}}\rangle/\sqrt{p_\theta^{\text{ps}}}$  is substituted into Eq. 3, the derivatives of  $p_\theta^{\text{ps}}$  cancel, such that

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4 \langle \dot{\psi}_\theta^{\text{ps}} | \dot{\psi}_\theta^{\text{ps}} \rangle \frac{1}{p_\theta^{\text{ps}}} - 4 |\langle \dot{\psi}_\theta^{\text{ps}} | \psi_\theta^{\text{ps}} \rangle|^2 \frac{1}{(p_\theta^{\text{ps}})^2}. \quad (5)$$

Equation 13 gives the quantum Fisher information available from a quantum state after its postselection. Unsurprisingly,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can exceed  $\mathcal{I}_Q(\theta|\hat{\rho}_\theta)$ , since  $p_\theta^{\text{ps}} \leq 1$ . Also classical systems can achieve such postselected information amplification (see Fig. 1). Unlike in the classical case, however,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can also exceed the Fisher information of an optimized prepare-measure experiment,  $(\Delta a)^2$ . We show how below.

### Quasiprobability representation

In classical mechanics, our knowledge of a point particle can be described by a probability distribution for the particle's position,  $\vec{x}$ , and momentum,  $\vec{k}$ :  $p(\vec{x}, \vec{k})$ . In quantum mechanics, position and momentum do not commute, and a state cannot generally be represented by a joint probability distribution over observables' eigenvalues. A quantum state can, however, be represented by a quasiprobability distribution. Many classes of quasiprobability distributions exist. The most famous is the Wigner function [45–47]. Such a distribution satisfies some, but not all, of Kolmogorov's axioms for probability

distributions [48]: the entries sum to unity, and marginalizing over the eigenvalues of every observable except one yields a probability distribution over the remaining observable's eigenvalues. A quasiprobability distribution can, however, have negative or nonreal values. Such values signal nonclassical physics in, for example, quantum computing and quantum chaos [2, 6, 49–57].

A cousin of the Wigner function is the Kirkwood-Dirac quasiprobability distribution [4–6]. This distribution, which has been referred to by several names across the literature, resembles the Wigner function for continuous systems. Unlike the Wigner functions, however, the Kirkwood-Dirac distribution is well-defined for discrete systems, even qubits. The Kirkwood-Dirac distribution has been used in the study of weak-value amplification [6, 52, 58–61], information scrambling [6, 56, 57, 62] and direct measurements of quantum wavefunctions [63–66]. Moreover, negative and nonreal values of the distribution have been linked to nonclassical phenomena [6, 52, 56, 57]. We cast the quantum Fisher information for a postselected prepare-measure experiment in terms of a *doubly extended*<sup>2</sup> Kirkwood-Dirac quasiprobability distribution [6]. We employ this distribution due to its usefulness as a mathematical tool: This distribution enables the proof that, in the presence of noncommuting observables, postselection can give a metrological protocol a nonclassical advantage.

Our distribution is defined in terms of eigenbases of  $\hat{A}$  and  $\hat{F}$ . Other quasiprobability distributions are defined in terms of bases independent of the experiment. For example, the Wigner function is often defined in the bases of the quadrature of the electric field or the position and momentum bases. However, basis-independent distributions can be problematic in the hunt for nonclassicality [2, 53]. Careful application, here, of the extended Kirkwood-Dirac distribution ties its nonclassical values to the operational specifics of the experiment.

To begin, we define the quasiprobability distribution of an arbitrary quantum state  $\hat{\rho}$ :

$$q_{a,a',f}^{\hat{\rho}} \equiv \langle f|a\rangle \langle a|\hat{\rho}|a'\rangle \langle a'|f\rangle. \quad (6)$$

Here,  $\{|f\rangle\}$ ,  $\{|a\rangle\}$  and  $\{|a'\rangle\}$  are bases for the Hilbert space on which  $\hat{\rho}$  is defined. We can expand  $\hat{\rho}$  [63, 64] as<sup>3</sup>

$$\hat{\rho} = \sum_{a,a',f} \frac{|a\rangle\langle f|}{\langle f|a\rangle} q_{a,a',f}^{\hat{\rho}}. \quad (7)$$

<sup>2</sup> The modifier “doubly extended” comes from the experiment in which one would measure the distribution: One would prepare  $\hat{\rho}$ , sequentially measure two observables weakly, and measure one observable strongly. The number of weak measurements equals the degree of the extension [6].

<sup>3</sup> If any  $\langle f|a\rangle = 0$ , we perturb one of the bases infinitesimally while preserving its orthonormality.

Let  $\{|a\rangle\} = \{|a'\rangle\}$  denote an eigenbasis of  $\hat{A}$ , and let  $\{|f\rangle\}$  denote an eigenbasis of  $\hat{F}$ . The reason for introducing a doubly extended distribution, instead of the standard Kirkwood-Dirac distribution  $q_{a,f}^{\hat{\rho}} \equiv \langle f|a\rangle \langle a|\hat{\rho}|f\rangle$ , is that  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can be expressed most concisely, naturally, and physically meaningfully in terms of  $q_{a,a',f}^{\hat{\rho}}$ . Later, we shall see how the nonclassical entries in  $q_{a,a',f}^{\hat{\rho}}$  and  $q_{a,f}^{\hat{\rho}}$  are related. We now express the postselected quantum Fisher information (Eq. 13) in terms of the quasiprobability values  $q_{a,a',f}^{\hat{\rho}_\theta}$  (Supp. Inf. 1).

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4 \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} aa' - 4 \left| \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a \right|^2, \quad (8)$$

where  $a$  and  $a'$  denote the eigenvalues associated with  $|a\rangle$  and  $|a'\rangle$ , respectively.<sup>4</sup> Equation 14 contains a conditional quasiprobability distribution,  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$ . If  $\hat{A}$  commutes with  $\hat{F}$ , as they do classically, then they share an eigenbasis for which  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}} \in [0, 1]$ , and the postselected quantum Fisher information is bounded as  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \leq (\Delta a)^2$ :

**Theorem 1** *In a classically commuting theory, no postselected prepare-measure experiment can generate more Fisher information than the optimized prepare-measure experiment.*

*Proof of Theorem 1.*—We upper-bound the right-hand side of Eq. 14. First, if  $\{|a\rangle\} = \{|a'\rangle\} = \{|f\rangle\}$  is a eigenbasis shared by  $\hat{A}$  and  $\hat{F}$ , Eq. 6 simplifies to a probability distribution:

$$q_{a,a',f}^{\hat{\rho}_\theta} = \langle a|\hat{\rho}_\theta|a'\rangle [ |f\rangle = |a\rangle ][ |a'\rangle = |f\rangle ] \in [0, 1], \quad (9)$$

where  $[X]$  is the Iverson bracket, which equals 1 if  $X$  is true and equals 0 otherwise. Second, summing  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$  over  $f \in \mathcal{F}^{\text{ps}}$ , we find

$$\sum_{f \in \mathcal{F}^{\text{ps}}} q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}} = \langle a|\hat{\rho}_\theta|a'\rangle \langle a'|\hat{F}|a\rangle/p_\theta^{\text{ps}}. \quad (10)$$

By the eigenbasis shared by  $\hat{A}$  and  $\hat{F}$ , the sum simplifies to  $\langle a|\hat{\rho}_\theta\hat{F}|a'\rangle [ |a'\rangle = |a\rangle ]/p_\theta^{\text{ps}}$ . We can thus rewrite Eq. 14:

$$\begin{aligned} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = & 4 \sum_{a,a'} \frac{\langle a|\hat{\rho}_\theta\hat{F}|a'\rangle [ |a'\rangle = |a\rangle ]}{p_\theta^{\text{ps}}} aa' \\ & - 4 \left| \sum_{a,a'} \frac{\langle a|\hat{\rho}_\theta\hat{F}|a'\rangle [ |a'\rangle = |a\rangle ]}{p_\theta^{\text{ps}}} a \right|^2 \end{aligned}$$

<sup>4</sup> We have suppressed degeneracy parameters  $\gamma$  in our notation for the states, e.g.,  $|a, \gamma\rangle \equiv |a\rangle$ .

$$= 4 \sum_a q_a a^2 - 4 \left( \sum_a q_a a \right)^2, \quad (11)$$

where we have defined the probabilities  $q_a \equiv \langle a|\hat{\rho}_\theta\hat{F}|a\rangle/p_\theta^{\text{ps}} = \sum_{f \in \mathcal{F}^{\text{ps}}} \langle a|\hat{\rho}_\theta|a\rangle [ |f\rangle = |a\rangle ]/p_\theta^{\text{ps}}$ .

Apart from the multiplicative factor of 4, Eq. 11 is in the form of a variance with respect to the observable's eigenvalues  $a$ . Thus, Eq. 11 is maximized when  $q_{a_{\min}} = q_{a_{\max}} = \frac{1}{2}$ :

$$\max_{\{q_a\}} \{\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})\} = (\Delta a)^2. \quad (12)$$

This Fisher-information bound must be independent of our choice of eigenbases of  $\hat{A}$  and  $\hat{F}$ . In summary, if  $\hat{A}$  commutes with  $\hat{F}$ , then all  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$  can be expressed as real and nonnegative, and  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \leq (\Delta a)^2$ .  $\square$

In contrast, if the quasiprobability distribution contains negative values, the postselected quantum Fisher information can violate the bound:  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) > (\Delta a)^2$ . In Supp. Inf. 2, we prove a second theorem:<sup>5</sup>

**Theorem 2** *An anomalous postselected Fisher information implies that the quantum Fisher information cannot decompose in terms of a nonnegative doubly extended Kirkwood-Dirac quasiprobability distribution.*

*Proof:* see Supplementary Note 2 for a proof.

This inability to decompose implies that  $\hat{A}$  fails to commute with  $\hat{F}$ . However, pairwise noncommutation of  $\hat{\rho}_\theta$ ,  $\hat{A}$  and  $\hat{F}$  is insufficient to enable anomalous values of  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ . For example, noncommutation could lead to a nonreal Kirkwood-Dirac distribution without any negative real components. Such a distribution cannot improve  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  beyond classical values. Furthermore, the presence or absence of commutation is a binary measure. In contrast, how much postselection improves  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  depends on how much negativity  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$  has. We build on this observation, and propose two experiments that yield anomalous Fisher-information values, in Supp. Infs. 3 and 4.<sup>6</sup>

As promised, we now address the relation between nonclassical entries in  $q_{a,a',f}^{\hat{\rho}_\theta}$  and nonclassical entries in  $q_{a,f}^{\hat{\rho}_\theta}$ . For pure states  $\hat{\rho} = |\Psi\rangle \langle\Psi|$ , the doubly extended quasiprobability distribution can be expressed time symmetrically in terms of the standard Kirkwood-Dirac distribution [4–6, 55–57]:  $q_{a,a',f}^{\hat{\rho}_\theta} = \frac{1}{p_f} q_{a,f}^{\hat{\rho}_\theta} (q_{a',f}^{\hat{\rho}_\theta})^*$ , where  $q_{a,f}^{\hat{\rho}_\theta} = \langle f|a\rangle \langle a|\hat{\rho}|f\rangle$  and  $p_f \equiv |\langle f|\Psi\rangle|^2$ .<sup>7</sup> Therefore, a negative  $q_{a,a',f}^{\hat{\rho}_\theta}$  implies negative or nonreal values of  $q_{a,f}^{\hat{\rho}_\theta}$ . Similarly, a negative  $q_{a,a',f}^{\hat{\rho}_\theta}$  implies a negative or nonreal

<sup>5</sup> The theorem's converse is not generally true.

<sup>6</sup> It remains an open question to investigate the relationship between Kirkwood-Dirac negativity in other metrology protocols with noncommuting operators, e.g., [67].

<sup>7</sup> See [20, 68] for discussions about time-symmetric interpretations of quantum mechanics.

weak value  $\langle f|a\rangle\langle a|\Psi\rangle/\langle f|\Psi\rangle$  [17], which possesses interesting ontological features (see below). Thus, an anomalous Fisher information is closely related to a negative or nonreal weak value. Had we weakly measured the observable  $|a\rangle\langle a|$  of  $\hat{\rho}_\theta$  with a qubit or Gaussian pointer particle before the postselection, and had we used a fine-grained postselection  $\{\hat{1} - \hat{F}, |f\rangle\langle f| : f \in \mathcal{F}^{\text{ps}}\}$ , the weak measurement would have yielded a weak value outside the eigenspectrum of  $|a\rangle\langle a|$ . It has been shown that such an anomalous weak value proves that quantum mechanics, unlike classical mechanics, is contextual: quantum outcome probabilities can depend on more than a unique set of underlying physical states [24, 36, 69]. If  $\hat{\rho}_\theta$  had undergone the aforementioned weak measurement, instead of the postselected prepare-measure experiment, the weak measurement's result would have signaled quantum contextuality. Consequently, a counterfactual connects an anomalous Fisher information and quantum contextuality. While counterfactuals create no problems in classical physics, they can lead to logical paradoxes in quantum mechanics [69–72]. Hence our counterfactual's implication for the ontological relation between an anomalous Fisher information and contextuality offers an opportunity for future investigation.

### Improved metrology via postselection

In every real experiment, the preparation and final measurement have costs, which we denote  $\mathcal{C}_P$  and  $\mathcal{C}_M$ , respectively. For example, a particle-number detector's *dead time*, the time needed to reset after a detection, associates a temporal cost with measurements [73]. Reference [38] concerns a two-level atom in a noisy environment. Liuzzo *et al.* detail the tradeoff between frequency estimation's time and energy costs. Standard quantum-metrology techniques, they show, do not necessarily improve metrology, if the experiment's energy is capped. Also, the cost of postprocessing can be incorporated into  $\mathcal{C}_M$ .<sup>8</sup> We define the information-cost rate as  $R(\theta) := N\mathcal{I}(\theta)/(N\mathcal{C}_P + N\mathcal{C}_M) = \mathcal{I}(\theta)/(\mathcal{C}_P + \mathcal{C}_M)$ . If our experiment conditions the execution of the final measurement on successful postselection of a fraction  $p_\theta^{\text{ps}}$  of the states, we include a cost of postselection,  $\mathcal{C}_{\text{ps}}$ . We define the postselected experiment's information-cost rate as  $R^{\text{ps}}(\theta) := Np_\theta^{\text{ps}}\mathcal{I}^{\text{ps}}(\theta)/(N\mathcal{C}_P + N\mathcal{C}_{\text{ps}} + Np_\theta^{\text{ps}}\mathcal{C}_M) = p_\theta^{\text{ps}}\mathcal{I}^{\text{ps}}(\theta)/(\mathcal{C}_P + \mathcal{C}_{\text{ps}} + p_\theta^{\text{ps}}\mathcal{C}_M)$ , where  $\mathcal{I}^{\text{ps}}(\theta)$  is the Fisher information conditioned on successful postselection. Generalizing the following arguments to preparation and measurement costs that differ between the postselected and nonpostselected experiments is straightforward.

<sup>8</sup> In an experiment, these costs can be multivariate functions that reflect the resources and constraints. Such a function could combine a detector's dead time with the monetary cost of liquid helium and a graduate student's salary. However, presenting the costs in a general form benefits this platform-independent work.

In classical experiments, postselection can improve the information-cost rate. See Fig. 1 for an example. But can postselection improve the information-cost rate in a classical experiment with information-optimized inputs? Theorem 1 answered this question in the negative.  $\mathcal{I}^{\text{ps}}(\theta) \leq \max\{\mathcal{I}(\theta)\}$  in every classical experiment. The maximization is over all physically accessible inputs and final measurements. A direct implication is that  $R^{\text{ps}}(\theta) \leq \max\{R(\theta)\}$ .

In quantum mechanics,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can exceed  $\max_{\hat{\rho}_\theta}\{\mathcal{I}_Q(\theta|\hat{\rho}_\theta)\} = (\Delta a)^2$ . This result would be impossible classically. Anomalous Fisher-information values require quantum negativity in the doubly extended Kirkwood-Dirac distribution. Consequently, even compared to quantum experiments with optimized input states, postselection can raise information-cost rates beyond classically possible rates:  $R^{\text{ps}}(\theta) > \max\{R(\theta)\}$ . This result generalizes the metrological advantages observed in the measurements of weak couplings, which also require noncommuting operators. References [74–83] concern metrology that involves weak measurements of the following form. The primary system S and the pointer P begin in a pure product state  $|\Psi_S\rangle \otimes |\Psi_P\rangle$ ; the coupling Hamiltonian is a product  $\hat{H} = \hat{A}_S \otimes \hat{A}_P$ ; the unknown coupling strength  $\theta$  is small; and just the system is postselected. Our results govern arbitrary input states, arbitrary Hamiltonians (that satisfy Stone's theorem), arbitrarily large coupling strengths  $\theta$ , and arbitrary projective postselections. Our result shows that postselection can improve quantum parameter estimation in experiments where the final measurement's cost outweighs the combined costs of state preparation and postselection:  $\mathcal{C}_M \gg \mathcal{C}_P + \mathcal{C}_{\text{ps}}$ . Earlier works identified that the Fisher information from nonrenormalized trials that succeed in the postselection cannot exceed the Fisher information averaged over all trials, including the trials in which the postselection fails [84, 85].<sup>9</sup> In accordance with practical metrology, not only the Fisher information, but also measurements' experimental costs, underlie our results.

So far, we have shown that  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can exceed  $(\Delta a)^2$ . But how large can  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  grow? In Supp. Inf. 3, we show that, if the generator  $\hat{A}$  has  $M \geq 3$  not-all-identical eigenvalues, there is no upper bound on  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ . If  $\mathcal{C}_P$  and  $\mathcal{C}_{\text{ps}}$  are negligible compared to  $\mathcal{C}_M$ , then there is no theoretical cap on how large  $R^{\text{ps}}(\theta)$  can grow. In general, when  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \rightarrow \infty$ ,  $p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) < (\Delta a)^2$ , such that information is lost

<sup>9</sup> Reference [43] considered squeezed coherent states as metrological probes in specific weak-measurement experiments. It is shown that postselection can improve the signal-to-noise ratio, irrespectively of whether the analysis includes the failed trials. However, this work concerned nonpostselected experiments in which only the probe state was measured. Had it been possible to successfully measure also the target system, the advantage would have disappeared.

in the events discarded by postselection. But if  $\hat{A}$  has doubly degenerate minimum and maximum eigenvalues,  $p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can approach  $(\Delta a)^2$  while  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  approaches infinity (see Supp. Inf. 4). In such a scenario, postselection can improve information-cost rates, as long as  $\mathcal{C}_{\text{ps}} < (1 - p_\theta^{\text{ps}})\mathcal{C}_M$ —a significantly weaker requirement than  $\mathcal{C}_M \gg \mathcal{C}_P + \mathcal{C}_{\text{ps}}$ .

## DISCUSSION

From a practical perspective, our results highlight an important quantum asset for parameter-estimation experiments with expensive final measurements. In some scenarios, the postselection’s costs exceed the final measurement’s costs, as an unsuccessful postselection might require fast feedforward to block the final measurement. But in single-particle experiments, the postselection can be virtually free and, indeed, unavoidable: an unsuccessful postselection can destroy the particle, precluding the triggering of the final measurement’s detection apparatus [86]. Thus, current single-particle metrology could benefit from postselected improvements of the Fisher information. A photonic experimental test of our results is currently under investigation.

From a fundamental perspective, our results highlight the strangeness of quantum mechanics as a noncommuting theory. Classically, an increase of the Fisher information via postselection can be understood as the *a posteriori* selection of a better input distribution. But it

is nonintuitive that quantum mechanical postselection can enable a quantum state to carry more Fisher information than the best possible input state could. The optimized Cramér-Rao bound, obtained from Eq. 4, can be written in the form of an uncertainty relation:  $\sqrt{\text{Var}(\theta_e)}(\Delta a) \geq 1$  [7]. Our results highlight the probabilistic possibility of violating this bound. More generally, the information-cost rate’s ability to violate a classical bound leverages negativity, a nonclassical resource in quantum foundations, for metrological advantage.

## Acknowledgements

The authors would like to thank Justin Dressel, Nicolas Delfosse, Matthew Pusey, Noah Lupu-Gladstein, Aharon Brodutch and Jan-Åke Larsson for useful discussions. D.R.M.A.-S. acknowledges support from the EPSRC, the Sweden-America Foundation, Hitachi Ltd, the Lars Hierta Memorial Foundation and Girton College. N.Y.H. was supported by an NSF grant for the Institute for Theoretical Atomic, Molecular, and Optical Physics at Harvard University and the Smithsonian Astrophysical Observatory. H.V.L. received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642688. A.A.L. acknowledges support from the EPSRC and Hitachi Ltd. S.L. was supported by NSF, AFOSR, and by ARO under the Blue Sky Initiative.

**SUPPLEMENTARY NOTE 1 – EXPRESSING THE POSTSELECTED QUANTUM FISHER INFORMATION IN TERMS OF THE KD DISTRIBUTION**

As shown in the Results section of our main paper, the postselected quantum Fisher information is given by

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4 \langle \dot{\psi}_\theta^{\text{ps}} | \dot{\psi}_\theta^{\text{ps}} \rangle \frac{1}{p_\theta^{\text{ps}}} - 4 \left| \langle \dot{\psi}_\theta^{\text{ps}} | \psi_\theta^{\text{ps}} \rangle \right|^2 \frac{1}{(p_\theta^{\text{ps}})^2}, \quad (13)$$

where nonrenormalized postselected quantum state is  $|\psi_\theta^{\text{ps}}\rangle = \hat{F}\hat{U}(\theta)|\Psi_0\rangle$ , where  $|\Psi_0\rangle\langle\Psi_0| \equiv \hat{\rho}_0$ .  $p_\theta^{\text{ps}} = \text{Tr}(\hat{F}\hat{\rho}_\theta)$  is the probability of postselection.

In this supplementary note, we show that Eq. 13 can be expressed in terms of the double-extended KD distribution:

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4 \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} aa' - 4 \left| \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a \right|^2, \quad (14)$$

The first term of the quantum Fisher information (Eq. 13) is

$$\frac{4}{p_\theta^{\text{ps}}} \langle \dot{\psi}_\theta^{\text{ps}} | \dot{\psi}_\theta^{\text{ps}} \rangle = \frac{4}{p_\theta^{\text{ps}}} \text{Tr}(\hat{F}\dot{U}(\theta)\hat{\rho}_0\dot{U}^\dagger(\theta)\hat{F}^\dagger) = \frac{4}{p_\theta^{\text{ps}}} \text{Tr}(\hat{F}\hat{A}\hat{\rho}_\theta\hat{A}) \quad (15)$$

$$= \frac{4}{p_\theta^{\text{ps}}} \text{Tr}\left(\sum_a |a\rangle\langle a| a \hat{\rho}_\theta \sum_{a'} |a'\rangle\langle a'| a' \sum_{f \in \mathcal{F}^{\text{ps}}} |f\rangle\langle f|\right), \quad (16)$$

where, in Eq. 16, we have expressed  $\hat{A}$  and  $\hat{F}$  in their corresponding eigendecompositions. This expression can be rewritten in terms of the doubly extended Kirkwood-Dirac quasiprobability distribution ( $q_{a,a',f}^{\hat{\rho}_\theta} = \langle f|a\rangle\langle a|\hat{\rho}|a'\rangle\langle a'|f\rangle$ ):

$$\frac{4}{p_\theta^{\text{ps}}} \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \text{Tr}\left(aa' q_{a,a',f}^{\hat{\rho}_\theta} \frac{|a\rangle\langle f|}{\langle f|a\rangle}\right) = \frac{4}{p_\theta^{\text{ps}}} \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} q_{a,a',f}^{\hat{\rho}_\theta} aa'. \quad (17)$$

Similarly, the second term of Eq. 13 is

$$\frac{4}{(p_\theta^{\text{ps}})^2} \left| \langle \dot{\psi}_\theta^{\text{ps}} | \dot{\psi}_\theta^{\text{ps}} \rangle \right|^2 = \frac{4}{(p_\theta^{\text{ps}})^2} \left| \text{Tr}(\hat{F}\hat{\rho}_\theta\hat{A}) \right|^2 = \frac{4}{(p_\theta^{\text{ps}})^2} \left| \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} q_{a,a',f}^{\hat{\rho}_\theta} a \right|^2. \quad (18)$$

Combining the expressions above gives Eq. 14:

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4 \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} aa' - 4 \left| \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a \right|^2. \quad (19)$$

**SUPPLEMENTARY NOTE 2 – PROOF OF THEOREM 2**

Here, we prove Theorem 2. First, we upper-bound the right-hand side of Eq. 14, assuming that all  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}} \in [0, 1]$ . We label the  $M$  eigenvalues of  $\hat{A}$  and arrange them in increasing order:  $a_1, a_2, \dots, a_M$ , such that  $a_1 \equiv a_{\min}$  and  $a_M \equiv a_{\max}$ . Initially, we assume that the 0-point of the eigenvalue axis is set such that  $a_1 = 0$  and  $a_M = \Delta a$ . In this scenario, all the components of the first term of Eq. 14 are nonnegative. We temporarily ignore the form of  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$ , and treat this ratio as a general quasiprobability distribution. Then,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  maximizes when  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$  vanishes at all  $a'$  values except  $a' = a_{\max}$ . We define  $q_a \equiv \sum_{a',f \in \mathcal{F}^{\text{ps}}} q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$ , such that all  $q_a \in [0, 1]$  and  $\sum_a q_a = 1$ . If  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}}$  is nonzero only when  $a' = a_{\max}$ , Eq. 14 becomes

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4a_M \sum_a q_a a - 4 \left( \sum_a q_a a \right)^2. \quad (20)$$

Expanding each sum, we obtain

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = 4a_M(q_{a_1}a_1 + K + q_{a_M}a_M) - 4(q_{a_1}a_1 + K + q_{a_M}a_M)^2 \quad (21)$$

$$= 4a_M(K + q_{a_M}a_M) - 4(K + q_{a_M}a_M)^2, \quad (22)$$

where we used  $q_{a_1}a_1 = 0$  and defined  $K \equiv \sum_{a \in \{a_2, \dots, a_{M-1}\}} q_a a \leq a_M$ . As  $\hat{A}$  is not totally degenerate,  $a_M \neq 0$ , and Eq. 22 is maximized when  $q_{a_M} = (a_M - 2K)/(2a_M)$ . This yields

$$\max\{\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})\} = a_M^2 = (\Delta a)^2, \quad (23)$$

where we have recalled that  $a_M = \Delta a$ .

We are left with proving that we can always set  $a_1 = 0$  and  $a_M = \Delta a$ . We continue to assume that  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}} \in [0, 1]$ , and we shift all the eigenvalues by a constant real value  $\delta_a$ . The effect on  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  is

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \rightarrow 4 \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} (a + \delta_a)(a' + \delta_a) - 4 \left[ \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} (a + \delta_a) \right]^2 \quad (24)$$

$$= 4 \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} aa' - 4 \left[ \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a \right]^2 + 4\delta_a \left( \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a - \sum_{\substack{a,a', \\ f \in \mathcal{F}^{\text{ps}}}} \frac{q_{a,a',f}^{\hat{\rho}_\theta}}{p_\theta^{\text{ps}}} a' \right) = \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}). \quad (25)$$

The last equality holds because  $q_{a,a',f}^{\hat{\rho}_\theta} = (q_{a',a,f}^{\hat{\rho}_\theta})^*$  generally and we are assuming that  $q_{a,a',f}^{\hat{\rho}_\theta} \in \mathbb{R}$ . Consequently, if all  $q_{a,a',f}^{\hat{\rho}_\theta}/p_\theta^{\text{ps}} \in [0, 1]$ , then  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \leq (\Delta a)^2$ . The second term of Eq. 14 cannot be decreased by imaginary values in  $q_{a,a',f}^{\hat{\rho}_\theta}$ . Moreover, the first term is necessarily real and nonnegative. Thus imaginary elements  $q_{a,a',f}^{\hat{\rho}_\theta}$  cannot increase  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ . If  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) > (\Delta a)^2$ , then  $q_{a,a',f}^{\hat{\rho}_\theta}$  must have negative entries.  $\square$

### SUPPLEMENTARY NOTE 3 – INFINITE POSTSELECTED QUANTUM FISHER INFORMATION

Here, we show that the postselected quantum Fisher information  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can approach infinity. The proof is by example; other examples might exist.

We assume that the generator  $\hat{A}$  has  $M \geq 3$  eigenvalues that are not all identical. We also assume that we possess an estimate  $\theta_0$  that lies close to the true value of  $\theta$ :  $\delta_\theta \equiv \theta - \theta_0$ , with  $|\delta_\theta| \ll 1$ . (The derivation of the quantum Fisher information also rests on the assumption that one has access to such an estimate [13].)

By Eqs. 13, 15 and 18,

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \frac{4}{p_\theta^{\text{ps}}} \text{Tr}(\hat{F} \hat{A} \hat{U}(\theta) \hat{\rho}_0 \hat{U}(\theta)^\dagger \hat{A}) - \frac{4}{(p_\theta^{\text{ps}})^2} \left| \text{Tr}(\hat{F} \hat{U}(\theta) \hat{\rho}_0 \hat{U}(\theta)^\dagger \hat{A}) \right|^2. \quad (26)$$

We now choose  $\hat{F}$  and  $\hat{\rho}_0$  such that  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  approaches infinity. Crudely,  $p_\theta^{\text{ps}}$  must approach 0 while  $\text{Tr}(\hat{F} \hat{A} \hat{U}(\theta) \hat{\rho}_0 \hat{U}(\theta)^\dagger \hat{A})$  either stays constant or approaches 0 more slowly. We label the  $M$  eigenvalues of  $\hat{A}$  and arrange them in increasing order:  $a_1, a_2, \dots, a_M$ , such that  $a_1 \equiv a_{\min}$  and  $a_M \equiv a_{\max}$ .

First, we choose  $\hat{F} = |f_1\rangle\langle f_1| + |f_2\rangle\langle f_2|$ , where

$$|f_1\rangle \equiv \frac{|a_{\max}\rangle + |a_{\min}\rangle}{\sqrt{2}}, \quad (27)$$

$$|f_2\rangle \equiv \frac{\frac{i}{\sqrt{2}}(|a_{\max}\rangle - |a_{\min}\rangle) + |a_k\rangle}{\sqrt{2}}, \quad (28)$$

and  $|a_k\rangle \neq |a_{\max}\rangle, |a_{\min}\rangle$ . We also choose  $\hat{\rho}_0 = |\Psi_0\rangle\langle\Psi_0|$  such that

$$|\Psi_0\rangle \equiv |\Psi_0(\theta_0, \phi)\rangle = \hat{U}^\dagger(\theta_0) \frac{1}{\sqrt{2}} \left\{ [\cos(\phi) - \sin(\phi)] \frac{i}{\sqrt{2}} (|a_{\min}\rangle - |a_{\max}\rangle) + [\cos(\phi) + \sin(\phi)] |a_k\rangle \right\}. \quad (29)$$

$\phi \approx 0$  is a parameter that can be tuned to maximize the postselected Fisher information for a given approximation accuracy  $\delta_\theta$ . As  $\phi$  is a parameter of the input state, variations in the Fisher information with  $\phi$  will reflect the effects of disturbances to the input state. Substituting the expressions for  $\hat{F}$  and  $\hat{\rho}_0$  into Eq. 26, we find

$$\begin{aligned} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) &= 8\{5 - 2\cos(2\phi)(\cos[(a_M - a_k)\delta_\theta] + \cos[(a_k - a_1)\delta_\theta]) + \cos[(a_M - a_1)\delta_\theta][\sin(2\phi) - 1] - \sin(2\phi)\}^{-2} \\ &\times \left\{ 2a_M^2 - a_M a_k + a_k^2 + 2a_1^2 - (3a_M + a_k)a_1 + (a_M - a_k)(a_k - a_1)\cos(4\phi)(\cos[(a_M - a_1)\delta_\theta] - 1) \right. \\ &+ (a_M - a_k)(a_k - a_1)\cos[(a_M - a_1)\delta_\theta] + 2(a_M - a_1)\cos(2\phi)((a_1 - a_k)\cos[(a_M - a_k)\delta_\theta] \\ &+ (a_k - a_M)\cos[(a_k - a_1)\delta_\theta]) - 2(a_M - a_1)^2\sin(2\phi) + (a_M - a_1)((a_k - a_1)\cos[(a_M - a_k)\delta_\theta] \\ &\left. + (a_M - a_k)\cos[(a_k - a_1)\delta_\theta])\sin(4\phi) \right\}. \end{aligned} \quad (30)$$

The postselection probability is

$$p_\theta^{\text{ps}} = \frac{1}{8}\{5 - 2\cos(2\phi)(\cos[(a_M - a_k)\delta_\theta] + \cos[(a_k - a_1)\delta_\theta]) + \cos[(a_M - a_1)\delta_\theta][\sin(2\phi) - 1] - \sin(2\phi)\}. \quad (31)$$

In the limit as our estimate  $\theta_0$  approaches the true value of  $\theta$ , such that  $\delta_\theta \rightarrow 0$ ,

$$\lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} = \sin^2(\phi), \quad (32)$$

$$\lim_{\delta_\theta \rightarrow 0} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \frac{(\cot(\phi) - 1)^2}{2}(\Delta a)^2, \quad \text{and} \quad (33)$$

$$\lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \frac{1}{2}[1 - \sin(2\phi)](\Delta a)^2. \quad (34)$$

In the limit as  $\phi \rightarrow 0$ ,

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \right] = 0, \quad (35)$$

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \right] = \infty, \quad \text{and} \quad (36)$$

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \right] = \frac{1}{2}(\Delta a)^2. \quad (37)$$

According to Eq. 36, if first  $\delta_\theta$  and then  $\phi$  approaches 0 in Eq. 30,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  approaches infinity.

There are a few points to note. First,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  diverges in the two ordered limits. In any real experiment, one could not blindly set  $\phi = 0$ , but would have to choose  $\phi$  based on an estimate of  $\theta$ . Second, if  $\delta_\theta \approx 0$ , then  $\theta_0 \approx \theta$ , and the pre-experiment variance of our initial estimate  $\theta_0$ ,  $\text{Var}(\theta_0)$ , must be small. That is, we begin the experiment with much information about  $\theta$ . Guided by the Cramér-Rao bound, we expect that, in a useful experiment,  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  would grow large, while  $1/\text{Var}(\theta_0) < \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ . Figure 3 shows  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \times \text{Var}(\theta_0)$  as a function of  $\phi$  and  $\delta_\theta$  for an experiment where  $a_1 = -1$ ,  $a_k = 1$ ,  $a_M = 3$  and  $\text{Var}(\theta_0) = 10^{-6}$ . If  $\theta_0$  is within a few  $\sigma_{\theta_0} \equiv \sqrt{\text{Var}(\theta_0)}$  of  $\theta$ , then  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \times \text{Var}(\theta_0) \gg 1$ . Figure 3 shows that large values of  $1/\delta_\theta$  can result in even larger values of  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ . Figure 3 also illustrates the effect of input-state disturbances of  $\phi$  on  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \times \text{Var}(\theta_0)$ . Third, while the theoretical strategy investigated in this appendix achieves an infinite postselected quantum Fisher information, the postselection also ‘‘wastes’’ information as  $\lim_{\phi \rightarrow 0}[\lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})] < (\Delta a)^2$ . If  $\hat{A}$  possesses certain properties, it is possible to avoid wasting information through the postselection; we show how in the following appendix.

#### SUPPLEMENTARY NOTE 4 – INFINITE POSTSELECTED QUANTUM FISHER INFORMATION WITHOUT LOSS OF INFORMATION

If the generator  $\hat{A}$  has  $M \geq 4$  eigenvalues, and the minimum and maximum eigenvalues are both at least doubly degenerate, then  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  can approach infinity without information’s being lost in the events discarded by postselection. We show how below.

First, we assign the orthonormal eigenvectors  $|a_{\min_1}\rangle$  and  $|a_{\min_2}\rangle$  to the eigenvalues  $a_1 = a_{\min}$  and  $a_2 = a_{\min}$ , respectively. Here, we have reused the eigenvalue notation from Supp. Mat. . Similarly, we assign the orthonormal

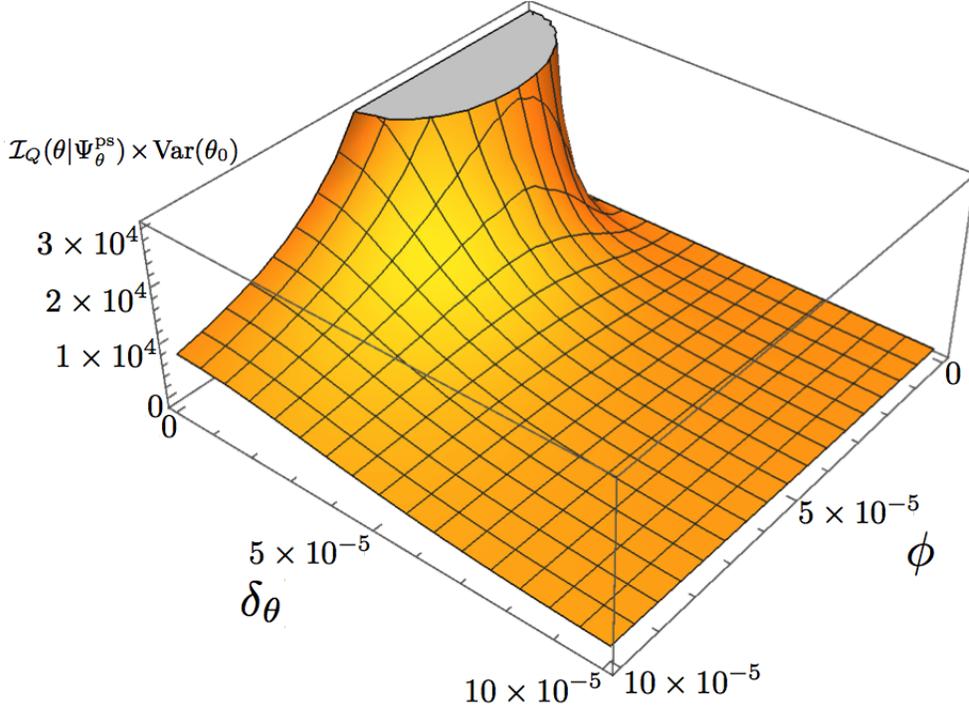


FIG. 3. **Scaled postselected quantum Fisher information.** The figure shows the postselected quantum Fisher information (Eq. 30) multiplied by the pre-experiment variance  $\text{Var}(\theta_0)$  as a function of  $\phi$  and  $\delta_\theta$ . For small values of  $\delta_\theta$  and  $\phi$ , the value of  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \times \text{Var}(\theta_0)$  diverges. The eigenvalues  $a_1$ ,  $a_k$  and  $a_M$  are set to  $-1$ ,  $1$  and  $3$ , respectively.  $\text{Var}(\theta_0)$  was set to  $1 \times 10^{-6}$ .

eigenvectors  $|a_{\max_1}\rangle$  and  $|a_{\max_2}\rangle$  to the eigenvalues  $a_M = a_{\max}$  and  $a_{M-1} = a_{\max}$ , respectively. Second, we set  $\hat{F} = |f_1\rangle\langle f_1| + |f_2\rangle\langle f_2|$ , where

$$|f_1\rangle \equiv \frac{|a_{\max_2}\rangle - |a_{\min_1}\rangle}{\sqrt{2}}, \quad (38)$$

$$|f_2\rangle \equiv \frac{|a_{\min_2}\rangle - |a_{\max_1}\rangle}{\sqrt{2}}. \quad (39)$$

We also choose  $|\Psi_0\rangle$  such that

$$|\Psi_0(\theta_0, \phi)\rangle = \hat{U}^\dagger(\theta_0) \frac{1}{2} \{ [\cos(\phi) - \sin(\phi)](|a_{\max_2}\rangle + |a_{\min_2}\rangle) + [\sin(\phi) + \cos(\phi)](|a_{\max_1}\rangle + |a_{\min_1}\rangle) \}. \quad (40)$$

As in App. ,  $\phi \approx 0$  is a parameter that can be tuned to maximize  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$  for a given approximation accuracy of  $\delta_\theta$ .

Substituting the expressions for  $\hat{F}$  and  $\hat{\rho}_0$  into Eq. 26, we find

$$\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \frac{\sin^2(2\phi)(a_M - a_1)^2}{(1 - \cos(2\phi) \cos[(a_M - a_1)\delta_\theta])^2}. \quad (41)$$

The postselection probability is

$$p_\theta^{\text{ps}} = \frac{1}{2} \{ 1 - \cos(2\phi) \cos[(a_M - a_1)\delta_\theta] \}. \quad (42)$$

Again, we investigate the limit as our estimate  $\theta_0$  approaches the true value of  $\theta$ :

$$\lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} = \sin^2(\phi), \quad (43)$$

$$\lim_{\delta_\theta \rightarrow 0} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \cot^2(\phi)(\Delta a)^2, \quad \text{and} \quad (44)$$

$$\lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = \cos^2(\phi)(\Delta a)^2. \quad (45)$$

In the limit as  $\phi \rightarrow 0$ ,

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \right] = 0, \quad (46)$$

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \right] = \infty, \quad \text{and} \quad (47)$$

$$\lim_{\phi \rightarrow 0} \left[ \lim_{\delta_\theta \rightarrow 0} p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) \right] = (\Delta a)^2. \quad (48)$$

In conclusion, the above strategy allows us to obtain an infinite value for  $\mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}})$ , while  $p_\theta^{\text{ps}} \times \mathcal{I}_Q(\theta|\Psi_\theta^{\text{ps}}) = (\Delta a)^2$ . No information is lost in the postselection. As in Supp Mat. , the results hold for the two ordered limits.

- [1] N. Lütkenhaus and S. M. Barnett, Phys. Rev. A **51**, 3340 (1995).
- [2] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).
- [3] C. Ferrie and J. Emerson, **41**, 352001 (2008).
- [4] J. G. Kirkwood, Phys. Rev. **44**, 31 (1933).
- [5] P. A. M. Dirac, Rev. Mod. Phys. **17**, 195 (1945).
- [6] N. Yunger Halpern, B. Swingle, and J. Dressel, Phys. Rev. A **97**, 042105 (2018).
- [7] V. Giovannetti, S. Lloyd, and L. Maccone, Nature photonics **5**, 222 (2011).
- [8] V. Giovannetti, S. Lloyd, and L. Maccone, Physical review letters **96**, 010401 (2006).
- [9] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, Phys. Rev. Lett. **107**, 080504 (2011).
- [10] R. Demkowicz-Dobrzański and L. Maccone, Phys. Rev. Lett. **113**, 250801 (2014).
- [11] L. Maccone, Phys. Rev. A **88**, 042109 (2013).
- [12] H. C. W., *Quantum Detection and Estimation Theory*, 1st ed. (Academic Press, New York, NY, USA, 1976) p. 320.
- [13] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (John Wiley and Sons Inc., Hoboken, New Jersey, USA, 2006).
- [15] H. Cramér, *Mathematical methods of statistics (PMS-9)*, Vol. 9 (Princeton University Press, 2016).
- [16] C. R. Rao, in *Breakthroughs in statistics* (Springer, 1992) pp. 235–247.
- [17] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).
- [18] M. S. Leifer and R. W. Spekkens, Phys. Rev. Lett. **95**, 200405 (2005).
- [19] J. Tollaksen, J. Phys. A **40**, 9033 (2007).
- [20] Y. Aharonov and L. Vaidman, *Time in quantum mechanics* (Springer, 2008) pp. 399–447.
- [21] J. Dressel, S. Agarwal, and A. N. Jordan, Phys. Rev. Lett. **104**, 240401 (2010).
- [22] L. Vaidman, Phys. Rev. A **87**, 052104 (2013).
- [23] C. Ferrie and J. Combes, Phys. Rev. Lett. **113**, 120404 (2014).
- [24] M. F. Pusey, Phys. Rev. Lett. **113**, 200401 (2014).
- [25] M. F. Pusey and M. S. Leifer, arXiv preprint arXiv:1506.07850 (2015).
- [26] D. R. M. Arvidsson-Shukur and C. H. W. Barnes, Phys. Rev. A **94**, 062303 (2016).
- [27] D. R. M. Arvidsson-Shukur, A. N. O. Gottfries, and C. H. W. Barnes, Phys. Rev. A **96**, 062316 (2017).
- [28] D. Schmid and R. W. Spekkens, Physical Review X **8**, 011015 (2018).
- [29] D. R. M. Arvidsson-Shukur and C. H. W. Barnes, Phys. Rev. A **99**, 060102 (2019).
- [30] V. Cimini, I. Gianani, F. Piacentini, I. P. Degiovanni, and M. Barbieri, **5**, 025007 (2020).
- [31] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, Phys. Rev. D **40**, 2112 (1989).
- [32] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Reviews of Modern Physics **86**, 307 (2014).
- [33] S. Pang, J. Dressel, and T. A. Brun, Phys. Rev. Lett. **113**, 030401 (2014).
- [34] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, Physical Review X **4**, 011031 (2014).
- [35] J. Harris, R. W. Boyd, and J. S. Lundeen, Phys. Rev. Lett. **118**, 070802 (2017).
- [36] R. Kunjwal, M. Lostaglio, and M. F. Pusey, arXiv preprint arXiv:1812.06940 (2018).
- [37] L. Xu, Z. Liu, A. Datta, G. C. Knee, J. S. Lundeen, Y.-q. Lu, and L. Zhang, arXiv preprint arXiv:2005.03629 (2020).
- [38] P. Liuzzo-Scorpo, L. A. Correa, F. A. Pollock, A. Górecka, K. Modi, and G. Adesso, New Journal of Physics **20**, 063009 (2018).
- [39] P. Lipka-Bartosik and R. Demkowicz-Dobrzański, Journal of Physics A: Mathematical and Theoretical **51**, 474001 (2018).
- [40] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).
- [41] D. Petz, *Quantum information theory and quantum statistics* (Springer Science & Business Media, 2007).
- [42] D. Petz and C. Ghinea, *Quantum Probability and Related Topics* (World Scientific, 2011) pp. 261–281.
- [43] S. Pang and T. A. Brun, Phys. Rev. Lett. **115**, 120401 (2015).

- [44] M. H. Stone, *Ann. Math* **33**, 643 (1932).
- [45] E. Wigner, *Phys. Rev.* **40**, 749 (1932).
- [46] W. K. Wootters, *Ann. Phys.* **176**, 1 (1987).
- [47] H. J. Carmichael, *Statistical methods in quantum optics 1 Master equations and Fokker-Planck equations* (Springer Science & Business Media, 2013).
- [48] A. N. Kolmogorov, *The Mathematical Gazette* **35** (1951), 10.2307/3611503.
- [49] C. Ferrie, *Rep. Prog. Phys.* **74**, 116001 (2011).
- [50] A. G. Kofman, S. Ashhab, and F. Nori, *Physics Reports* **520**, 43 (2012), nonperturbative theory of weak pre- and post-selected measurements.
- [51] M. Howard, J. Wallman, V. Veitch, and J. Emerson, *Nature* **510**, 351 (2014).
- [52] J. Dressel, *Phys. Rev. A* **91**, 032116 (2015).
- [53] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, *Phys. Rev. X* **5**, 021003 (2015).
- [54] N. Delfosse, C. Okay, J. Bermejo-Vega, D. E. Browne, and R. Raussendorf, *New Journal of Physics* **19**, 123024 (2017).
- [55] N. Yunger Halpern, *Phys. Rev. A* **95**, 012120 (2017).
- [56] N. Y. Halpern, A. Bartolotta, and J. Pollack, *Communications Physics* **2**, 1 (2019).
- [57] J. R. González Alonso, N. Yunger Halpern, and J. Dressel, *Phys. Rev. Lett.* **122**, 040404 (2019).
- [58] A. M. Steinberg, *Phys. Rev. A* **52**, 32 (1995).
- [59] L. M. Johansen, *Phys. Rev. A* **76**, 012119 (2007).
- [60] H. F. Hofmann, *New Journal of Physics* **14**, 043031 (2012).
- [61] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese, *Phys. Rev. Lett.* **117**, 170402 (2016).
- [62] R. Mohseninia, J. R. G. Alonso, and J. Dressel, *Phys. Rev. A* **100**, 062336 (2019).
- [63] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, *Nature* **474**, 188 (2011).
- [64] J. S. Lundeen and C. Bamber, *Phys. Rev. Lett.* **108**, 070402 (2012).
- [65] C. Bamber and J. S. Lundeen, *Phys. Rev. Lett.* **112**, 070405 (2014).
- [66] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen, *Phys. Rev. Lett.* **117**, 120401 (2016).
- [67] L. Sun, X. He, C. You, C. Lv, B. Li, S. Lloyd, and X. Wang, arXiv preprint arXiv:2004.01216 (2020).
- [68] M. S. Leifer and M. F. Pusey, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **473**, 20160607 (2017).
- [69] R. W. Spekkens, *Phys. Rev. A* **71**, 052108 (2005).
- [70] S. Kochen and E. P. Specker, in *The logico-algebraic approach to quantum mechanics* (Springer, 1975) pp. 293–328.
- [71] L. Hardy, *Phys. Rev. Lett.* **68**, 2981 (1992).
- [72] R. Penrose, *Shadows of the Mind: A Search for the Missing Science of Consciousness*, 1st ed. (Oxford University Press, Inc., New York, NY, USA, 1994) p. 240.
- [73] C. Greganti, P. Schiаны, I. A. Calafell, L. M. Procopio, L. A. Rozema, and P. Walther, *Opt. Express* **26**, 3286 (2018).
- [74] O. Hosten and P. Kwiat, *Science* **319**, 787 (2008).
- [75] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, *Phys. Rev. Lett.* **102**, 173601 (2009).
- [76] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, *Phys. Rev. A* **80**, 041803 (2009).
- [77] N. Brunner and C. Simon, *Phys. Rev. Lett.* **105**, 010405 (2010).
- [78] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, *Phys. Rev. A* **82**, 063822 (2010).
- [79] P. Egan and J. A. Stone, *Opt. Lett.* **37**, 4991 (2012).
- [80] H. F. Hofmann, M. E. Goggin, M. P. Almeida, and M. Barbieri, *Phys. Rev. A* **86**, 040102 (2012).
- [81] O. S. Magaña Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, *Phys. Rev. Lett.* **112**, 200401 (2014).
- [82] K. Lyons, J. Dressel, A. N. Jordan, J. C. Howell, and P. G. Kwiat, *Phys. Rev. Lett.* **114**, 170801 (2015).
- [83] J. Martínez-Rincón, C. A. Mullarkey, G. I. Viza, W.-T. Liu, and J. C. Howell, *Opt. Lett.* **42**, 2479 (2017).
- [84] C. Ferrie and J. Combes, *Phys. Rev. Lett.* **112**, 040406 (2014).
- [85] J. Combes, C. Ferrie, Z. Jiang, and C. M. Caves, *Phys. Rev. A* **89**, 052117 (2014).
- [86] I. A. Calafell, T. Strömberg, D. Arvidsson-Shukur, L. Rozema, V. Saggio, C. Greganti, N. Harris, M. Prabhu, J. Carolan, M. Hochberg, *et al.*, *npj Quantum Information* **5**, 61 (2019).