


Evading noise in multiparameter quantum metrology with indefinite causal order

Aaron Z. Goldberg  and Khabat Heshami 

National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario, Canada K1N 5A2
and Department of Physics, University of Ottawa, Advanced Research Complex, 25 Templeton Street, Ottawa, Ontario, Canada K1N 6N5

L. L. Sánchez-Soto 

Max-Planck-Institut für die Physik des Lichts, 91058 Erlangen, Germany
and Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain



(Received 23 March 2023; accepted 2 August 2023; published 20 September 2023)

Quantum theory allows the traversing of multiple channels in a superposition of different orders. When the order in which the channels are traversed is controlled by an auxiliary quantum system, various unknown parameters of the channels can be estimated by measuring only the control system, even when the state of the probe alone would be insensitive. Moreover, increasing the dimension of the control system increases the number of simultaneously estimable parameters, which has important metrological ramifications. We demonstrate this capability for simultaneously estimating both unitary and noise parameters, including multiple parameters from the same unitary such as rotation angles and axes and from noise channels such as depolarization, dephasing, and amplitude damping in arbitrary dimensions. We identify regimes of unlimited advantages, taking the form of p^2 smaller variances in estimation when the noise probability is $1 - p$, for both single and multiparameter estimations when using our schemes relative to any comparable scheme whose causal order is definite.

DOI: [10.1103/PhysRevResearch.5.033198](https://doi.org/10.1103/PhysRevResearch.5.033198)

I. INTRODUCTION

All measurements comprise four steps: Initializing a probe or receiver, letting the probe interact with some system whose properties are to be measured, performing a measurement on the probe by which to extract data, and estimating the unknown parameter based on the data [1]. Classical estimation theory dictates how to optimize the fourth step, quantum estimation theory the third, and judicious changes in the first can lead to remarkable advantages when using probes with particular quantum properties; the interaction step, in contradistinction, is typically taken to be immutable. Introducing indefinite causal order (ICO) provides a paradigm for changing the interaction step of a measurement protocol, thereby offering a further avenue for quantum advantages, which can now be exploited to great avail.

Quantum estimation theory establishes the potential advantages of quantum probe states and quantum measurement techniques for estimating parameters in a variety of physical processes [2–8]. This power has been demonstrated in remarkable experiments [9–13] and is expected to lead to practical, quantum-enhanced technologies in the near future [14–18]. The regime of multiparameter estimation is especially rich [19–30], with questions about incompatible observables [31–35], nuisance parameters [36,37], and

tradeoffs between parameters [38] rising to the fore, which is prominent because many practical measurement scenarios involve the simultaneous estimation of multiple parameters [39–46]. It is to this multiparameter scenario that we apply ICO to coax more practical advantages from quantum systems.

The idea of ICO stems from studies of causal structures in quantum gravity and quantum computation [47,48] and has since burgeoned into a pervasive research field. Incorporating ICO in particular tasks leads to enhancements relative to *quantum* advantages in computation [49–53], communication [54–60], cooling [61–64], work extraction [65–67], and sensing [68–72], many of which have been experimentally realized [73–83], along with more foundational ramifications [84–94]. Moreover, ICO can sometimes be used to inspire protocols with *definite* causal order (DCO) that outperform previously known methods [95]. The dramatic improvements possible in particular metrological tasks [70], as well as the ability to sense hitherto hidden parameters [96], motivate our current paper.

Quantum noise, in general, ruins many proposed quantum advantages [97–101], yet it is prevalent in all realistic scenarios across quantum information, making a quantum advantage in the presence of noise even more impressive. We recently showed that ICO confers dramatic advantages for estimating the phase of a unitary in arbitrary dimensions in the presence of depolarization noise, offering $O(p^2)$ smaller variances than any scheme with DCO when the depolarization strength is large and p is thereby small [96]. These results can even be obtained using maximally mixed probe states that are completely insensitive to such parameters when evolved with a DCO.

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

We here demonstrate how to apply ICO to a wide variety of estimation problems and find dramatic sensitivity advantages in the presence of noise for both single and multiparameter estimations. After first providing a background on ICO, its implementation using a quantum switch and a background on the quantum Fisher information (QFI) paradigm in Sec. II, we showcase our recent results for estimation of an arbitrarily large-dimensional unitary's phase in the presence of depolarization noise in Sec. III A, followed by advantages for dephasing (Sec. III B) and amplitude damping (Sec. III C) noise in arbitrary dimensions. The advantages, in terms of how much smaller the estimators' variances are for our ICO scheme relative to the best possible scheme with DCO, are on the order of $O[(\frac{p_A p_B}{p_A + p_B})^2]$ for depolarization, where small p_O means that channel O is very noisy; $O[(\frac{1}{2} - p_A)^2(\frac{1}{2} - p_B)^2]$ for dephasing along a particular axis, where small $|\frac{1}{2} - p_O|$ means that the dephasing (or spin-flip) channel is very noisy; and $O[\frac{p_A p_B}{(\sqrt{p_A + \sqrt{p_B}})^2 + c(\sqrt{p_A} - \sqrt{p_B})^2}]$ for amplitude damping noise with some dimension-dependent constant c , where small p_O again means that channel O is very noisy. All schemes have the amount of information decrease as the amount of noise increases, but ICO is more resilient to noise and therefore more efficient in terms of the number of times the unitary must be probed in the large-noise limit, as attested to by its advantageous scaling in noise parameters. Formally, *infinite* advantages are thus possible when either $p_A = 0$ or $p_B = 0$ for depolarization and amplitude damping, as well as when either or both of $p_A = \frac{1}{2}$ and $p_B = \frac{1}{2}$ for dephasing. By *infinite advantages*, we herein mean that ICO confers the ability to measure something that would be impossible without ICO. ICO is beneficial for metrology in and around these limits of when schemes with DCO fail or begin to fail due to being overwhelmed by noise.

In Sec. IV, we take the opportunity to show how ICO can be used, not just in the presence of noise but to characterize properties of the noise itself by developing the theory of ICO for multiparameter estimation. Noise characterization is paramount for developing quantum devices and quantum networks, in both practical and adversarial scenarios. We show there how ICO can be used to simultaneously measure parameters from the noisy channels and the unitary operator being applied, investigating all three noise scenarios in turn. The crucial upgrade required to be sensitive to more parameters is to increase the dimension of the control system governing the causal order of operations, which requires the ability to consider all orders of the unitary and noise channels. For simultaneously estimating the unitary's phase as well as the strengths of the two noise channels, both depolarization and dephasing again offer formally infinite advantages for ICO when one of the noise channels is completely depolarizing or completely dephasing. Depolarization channels with complete control of the order of the channels even allow estimation of the unitary's phase when *both* channels are completely depolarizing ($p_A = p_B = 0$) and, for amplitude damping channels, we qualitatively show ICO's advantage to rapidly grow with decreasing p_O . The amplitude damping channel can also be used with a higher-dimensional control to simultaneously estimate the unitary's phase and rotation axis in addition to noise parameters.

The aforementioned sections allow ICO to change the order in which the noise and unitary channels are applied on a probe state. In contrast, in all other studies of noisy metrology with ICO, multiple copies of the noisy unitary are applied in an indefinite order, with the causal relationship between the noise and unitary fixed in each channel [68,71,72,102–104]. Even when those studies find advantages for ICO, they tend to be small, as the crucial component of our work is controlling the very order in which the noise and unitaries are applied, even with a single copy of the unitary. For completeness, we show in Sec. V how multiple copies of the same unitary subject to the same depolarization channel, with a fixed relationship between the noise and unitary, can be augmented with ICO to simultaneously measure the unitary and noise parameters; this extends the lines of previous work to multiparameter estimation and arbitrary dimensional probe systems. We also discuss advantages in scaling for these related scenarios, showing how to decrease the variance in estimating a unitary's phase decreases by $O(p^{D-1})$ for D identical copies of the noisy unitary channel.

Finally, we observe in Sec. VI that, when the probe state is a qubit and is sent through arbitrary numbers of channels in arbitrary numbers of orders controlled by a control state with arbitrary dimensions, the measurements on the control are independent from the chosen probe state for a general class of channels. This, to our knowledge, is the second foray of ICO into the context of multiparameter metrology [104] and our widespread results indicate that ICO will remain a stalwart in this field.

II. BACKGROUND

A. Primer on indefinite causal order using quantum switches

When two independent quantum channels $\mathcal{E}^{(A)}$ and $\mathcal{E}^{(B)}$ act sequentially on a quantum state ρ_p with a DCO, the total evolution is governed by the sequential application:

$$\rho_p \mapsto \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}(\rho_p) \quad \text{or} \quad \rho_p \mapsto \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}(\rho_p). \quad (2.1)$$

We use ρ_p to denote the *probe* state. A quantum switch breaks from this paradigm by allowing an external quantum system to control the order in which two or more channels act on ρ_p . Such a device is sufficient for achieving a number of advantages in a variety of tasks and has been realized experimentally in groundbreaking experiments.

To wit, suppose the sequence is governed by the state of an auxiliary system, termed *control* state ρ_c . When ρ_c is in some state $|0\rangle\langle 0|$, the probe evolves following $\mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}(\rho_p)$, while $\rho_c = |1\rangle\langle 1|$ dictates the evolution $\mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}(\rho_p)$. What, then, occurs when ρ_c is prepared in a superposition of $|0\rangle$ and $|1\rangle$? This is the realm of ICO.

The dynamics are easiest to picture with unitary operations $\mathcal{E}^{(O)}(\bullet) = U^{(O)}(\bullet)U^{(O)\dagger}$. The total evolution is encapsulated by the unitary operator

$$\mathcal{U} = |0\rangle\langle 0| \otimes U^{(B)}U^{(A)} + |1\rangle\langle 1| \otimes U^{(A)}U^{(B)}, \quad (2.2)$$

acting on the joint state $\rho_c \otimes \rho_p$, which can be immediately verified for its action when ρ_c is in state $|0\rangle$ or $|1\rangle$. This leads to cross terms in the joint dynamics when the control

is prepared in some superposition state $\psi_0|0\rangle + \psi_1|1\rangle$:

$$\begin{aligned} \mathcal{U}\rho_c \otimes \rho_p \mathcal{U}^\dagger &= |\psi_0|^2|0\rangle\langle 0| \otimes \mathcal{E}^{(B)}[\mathcal{E}^{(A)}(\rho_p)] \\ &+ |\psi_1|^2|1\rangle\langle 1| \otimes \mathcal{E}^{(A)}[\mathcal{E}^{(B)}(\rho_p)] \\ &+ \psi_0\psi_1^*|0\rangle\langle 1| \otimes U^{(B)}U^{(A)}\rho_p U^{(B)\dagger}U^{(A)\dagger} \\ &+ (\psi_0\psi_1^*|0\rangle\langle 1| \otimes U^{(B)}U^{(A)}\rho_p U^{(B)\dagger}U^{(A)\dagger})^\dagger; \end{aligned} \tag{2.3}$$

the final two terms represent interference effects that have found a number of applications. A natural assumption throughout this paper is that none of the channels [here neither $U^{(A)}$ nor $U^{(B)}$] change on timescales relevant to the amount of time it takes ρ_p to traverse them. Even though each unitary $U^{(A)}$ and $U^{(B)}$ appears twice in \mathcal{U} , the quantum switch ensures that each channel is only probed once. This can be seen by considering auxiliary flag degrees of freedom in quantum states $|0\rangle_{FA}$ and $|0\rangle_{FB}$ that transform as $|n\rangle_{FO} \mapsto |n+1\rangle_{FO}$ whenever $U^{(O)}$ is applied to the system; the flag degrees of freedom factor out after the application of the switch and are uniquely in the states $|1\rangle_{FA}$ and $|1\rangle_{FB}$.

Similar dynamics result from quantum channels that are not unitary. For example, we can consider maps characterized by Kraus operators:

$$\mathcal{E}^{(O)}(\bullet) = \sum_l K_l^{(O)}(\bullet)K_l^{(O)\dagger}, \quad \sum_l K_l^{(O)\dagger}K_l^{(O)} = \mathbb{1}. \tag{2.4}$$

The total evolution is then governed by a quantum channel with Kraus operators of the form [48,79,80]

$$\mathcal{K}_{ij} = |0\rangle\langle 0| \otimes K_i^{(B)}K_j^{(A)} + |1\rangle\langle 1| \otimes K_j^{(A)}K_i^{(B)}, \tag{2.5}$$

acting on the joint state $\rho_c \otimes \rho_p$, which can again be immediately verified for its action when ρ_c is in state $|0\rangle$ or $|1\rangle$ and requires no correlations between the Kraus operators for distinct modes. This again leads to interference terms in the dynamics, which again can be seen when the control is prepared in the superposition state $\psi_0|0\rangle + \psi_1|1\rangle$:

$$\begin{aligned} &\sum_{i,j} \mathcal{K}_{ij}\rho_c \otimes \rho_p \mathcal{K}_{ij}^\dagger \\ &= |\psi_0|^2|0\rangle\langle 0| \otimes \mathcal{E}^{(B)}[\mathcal{E}^{(A)}(\rho_p)] \\ &+ |\psi_1|^2|1\rangle\langle 1| \otimes \mathcal{E}^{(A)}[\mathcal{E}^{(B)}(\rho_p)] \\ &+ \psi_0\psi_1^*|0\rangle\langle 1| \otimes \sum_{i,j} K_i^{(B)}K_j^{(A)}\rho_p K_i^{(B)\dagger}K_j^{(A)\dagger} \\ &+ \left(\psi_0\psi_1^*|0\rangle\langle 1| \otimes \sum_{i,j} K_i^{(B)}K_j^{(A)}\rho_p K_i^{(B)\dagger}K_j^{(A)\dagger} \right)^\dagger. \end{aligned} \tag{2.6}$$

The quantum-channel evolutions under ICO may be deduced by interpreting Kraus operators as remnants from unitary operations on an enlarged Hilbert space that have had the auxiliary degrees of freedom traced out. We can always consider the Kraus operators $K_i^{(O)}$ to represent the actions of unitary operators $U^{(O,O')}$ acting on $\rho_c \otimes |0\rangle_{O'}\langle 0|$ via

$$K_i^{(O)} = {}_{O'}\langle i|U^{(O,O')}|0\rangle_{O'}. \tag{2.7}$$

Assuming each of the sequential operations to possess their own auxiliary modes, we can enlarge the unitary operators of Eq. (2.2) to become

$$\mathcal{U} = |0\rangle\langle 0| \otimes U^{(B,B')}U^{(A,A')} + |1\rangle\langle 1| \otimes U^{(A,A')}U^{(B,B')}. \tag{2.8}$$

Tracing out the auxiliary modes from the evolution

$$\begin{aligned} &\rho_c \otimes \rho_p \otimes |0\rangle_{A'}\langle 0| \otimes |0\rangle_{B'}\langle 0| \\ &\mapsto \mathcal{U}\rho_c \otimes \rho_p \otimes |0\rangle_{A'}\langle 0| \otimes |0\rangle_{B'}\langle 0| \mathcal{U}^\dagger \end{aligned} \tag{2.9}$$

immediately yields the Kraus operators given by Eqs. (2.5) and (2.7). Notwithstanding this interpretation, only Kraus operators of the form of Eq. (2.5) reduce to the unitary \mathcal{U} in the limit of a single Kraus operator because a quantum switch is a superoperator that must act in the same manner regardless of the process in question [48]. In fact, any alternative Kraus-operator decompositions for the individual channels $\mathcal{E}^{(O)}$ will lead to the same overall dynamics when the alternative Kraus operators are fed into Eq. (2.5). As such, given only the two respective descriptions of channels $\mathcal{E}^{(A)}$ and $\mathcal{E}^{(B)}$, a quantum switch is guaranteed to lead to evolution with Kraus operators from Eq. (2.5) without requiring any control of the details of the channels or correlations between A and B .

This form of the Kraus operators arising from superpositions of sequences of operations holds true when there are arbitrary numbers of operations whose orders of application are being superposed. By increasing the dimension D of the control system, we can increase the number of possible orderings. If we label the Kraus operators from each channel A_j in the sequence by $K_i^{A_j}$, the control system can enable D different permutations of the channels A_j , leading to Kraus operators of the form

$$\mathcal{K}_{i_1, i_2, \dots, i_D} = \sum_{j=0}^{D-1} |j\rangle\langle j| \otimes K_{i_{\pi_j(0)}}^{(A_{\pi_j(0)})} K_{i_{\pi_j(1)}}^{(A_{\pi_j(1)})} \dots K_{i_{\pi_j(D-1)}}^{(A_{\pi_j(D-1)})}, \tag{2.10}$$

where we have denoted by $\pi_j(k)$ the k th element of the j th permutation of $(0, 1, \dots, D-1)$ and assumed there to be D channels without loss of generality [105]. Any time the control state is prepared in a superposition $\sum_j \psi_j|j\rangle$, interference terms with $j_1 \neq j_2$ will arise that can lead to unique effects in

$$\sum_{i_1 \dots i_D} \mathcal{K}_{i_1 \dots i_D} \rho_p \otimes \rho_c \mathcal{K}_{i_1 \dots i_D}^\dagger = \sum_{j_1 j_2} \psi_{j_1} \psi_{j_2}^* |j_1\rangle\langle j_2| \otimes \mathcal{R}_{j_1 j_2}. \tag{2.11}$$

Here,

$$\begin{aligned} \mathcal{R}_{j_1 j_2} &= \sum_{i_1, \dots, i_D} \left(K_{i_{\pi_{j_1}(0)}}^{(A_{\pi_{j_1}(0)})} \dots K_{i_{\pi_{j_1}(D-1)}}^{(A_{\pi_{j_1}(D-1)})} \right) \\ &\times \rho_p \left(K_{i_{\pi_{j_2}(0)}}^{(A_{\pi_{j_2}(0)})} \dots K_{i_{\pi_{j_2}(D-1)}}^{(A_{\pi_{j_2}(D-1)})} \right)^\dagger. \end{aligned} \tag{2.12}$$

In our paper, as is often the case, we will solely use properties of the control system to learn about the interactions of the probe. The control evolves to

$$\rho'_c = \sum_{j_1, j_2} \psi_{j_1} \psi_{j_2}^* |j_1\rangle\langle j_2| \mathcal{R}_{j_1 j_2}(\rho_p), \tag{2.13}$$

where we have defined the traces

$$R_{j_1 j_2} = \text{Tr}(\mathcal{R}_{j_1 j_2}). \tag{2.14}$$

Trace-preserving channels lead to $R_{jj}(\rho_p) = 1$; the interference terms with $R_{j_1 j_2}(\rho_p) < 1$ lead to entanglement between the control and the probe systems that can be used to estimate properties of the channels by measuring only the control. For consistency, we note that the case of two identical channels with a single Kraus operator $K^{(A)} = K^{(B)} = U$ simply has $R_{j_1 j_2} = \text{Tr}(\rho_p U^\dagger U) = 1$, such that the control only changes state when the channels A and B are nonunitary or not identical.

B. Quantum Fisher information

Suppose one has a set of parameters θ to estimate. Given access to a probability distribution $P(x|\theta)$ for some measurement with outcomes labeled by x , the Cramér-Rao bound dictates that the covariances between any estimators $\hat{\theta}_i$ of the parameters will locally be lower bounded by the inverse of the Fisher information (FI) matrix

$$\text{Cov}(\hat{\theta}_i, \hat{\theta}_j) \geq (\mathbf{F}_x^{-1}(\theta))_{ij}, \tag{2.15}$$

where the latter has components

$$[\mathbf{F}_x(\theta)]_{ij} = \sum_x P(x|\theta) \frac{\partial \ln P(x|\theta)}{\partial \theta_i} \frac{\partial \ln P(x|\theta)}{\partial \theta_j}. \tag{2.16}$$

Analogous expressions can be found for continuous measurement outcomes x with integrals replacing the sums. The QFI matrix provides the ultimate upper bound for \mathbf{F} for any given probe state and underlying values of θ , thereby providing the ultimate lower limit for the covariance matrix. Given a probe state that has evolved to depend on the parameters, ρ_θ , one can always define the symmetric logarithmic derivatives

$$\frac{\partial \rho_\theta}{\partial \theta_i} = \frac{\rho_\theta L_i + L_i \rho_\theta}{2} \tag{2.17}$$

to provide a matrix analog of the derivatives in Eq. (2.16), where L_i may depend on ρ_θ and θ and is always Hermitian. Then, the QFI matrix is defined componentwise as [20]

$$[\mathbf{Q}_{\rho_\theta}(\theta)]_{ij} = \frac{1}{2} \text{Tr}(\rho_\theta \{L_i, L_j\}), \tag{2.18}$$

where $\{ \cdot, \cdot \}$ stands for the anticommutator $\{A, B\} = AB + BA$. The matrix inequality

$$\mathbf{Q}_{\rho_\theta}(\theta) \geq \mathbf{F}_x(\theta) \tag{2.19}$$

always holds in the sense that $\mathbf{Q} - \mathbf{F}$ is always positive semidefinite. Remarkably, the most general probability distribution $P(x|\theta) = \text{Tr}(\Pi_x \rho_\theta)$ for a positive operator-valued measure (POVM) with elements $\{\Pi_x\}$ can always be optimized in the asymptotic limit, in the sense that, for any positive-definite weight matrix \mathbf{W} , there exists an optimal POVM such that

$$\text{Tr}[\mathbf{W} \text{Cov}(\hat{\theta}, \hat{\theta})] = f \text{Tr}[\mathbf{W} \mathbf{Q}_{\rho_\theta}^{-1}(\theta)]. \tag{2.20}$$

where $f = 1$ for single-parameter estimation and $1 \leq f \leq 2$ for multiparameter estimation and where the equality holds after many repeated optimal measurements [106,107]. This connects the ultimate lower bounds on the covariances of the

estimators to the ultimate measurement scheme for any probe state; the optimal overall protocol then involves optimizing the QFI matrix over all probe states. We are generous throughout with this factor of f : We use the QFI matrix for all schemes with fixed causal order, even though the results attainable will be smaller by a factor of f , and provide fixed measurement schemes for all of our new protocols with ICO, which can be directly fed into the Cramér-Rao bound of Eq. (2.16). This means that our quoted results hereafter for ICO may actually outperform schemes with a fixed causal order by an extra factor of f . Because the QFI matrix is additive when the same measurement process is repeated, we henceforth consider a single trial when comparing QFI values for different protocols.

C. Estimating the phase of a unitary

Consider any finite-dimensional unitary operator U , which can always be considered as an element of $\text{SU}(N)$ for some positive integer N without loss of generality. If the generators of $\text{SU}(N)$ are labeled by $\mathbf{G} = (G_1, G_2, \dots)$, then we can define $G_{\mathbf{n}} = \mathbf{n} \cdot \mathbf{G}$ for some unit vector \mathbf{n} and always express the unitary as

$$U(\theta, \mathbf{n}) = \exp(i\theta G_{\mathbf{n}}). \tag{2.21}$$

Estimating the phase θ is a basic problem with broad applications due to the ubiquity of unitary operations; we name interferometry, magnetometry, and imaging as examples. To fix the resources used in the estimation, we choose a particular irreducible representation of the Lie group with dimension d , equivalent to fixing the number of particles in or energy used by a probe state. Such a fixed irreducible representation has some eigenstates $|\pm \mathbf{n}\rangle$ of $G_{\mathbf{n}}$ with some maximal and minimal eigenvalues λ_{\pm} . Then, the best possible quantum strategy with DCO for estimating θ involves preparing the pure superposition state [108],

$$|\psi_{\text{opt}}\rangle = \frac{1}{\sqrt{2}}(|\mathbf{n}\rangle + |-\mathbf{n}\rangle), \tag{2.22}$$

and allowing it to evolve to $U|\psi\rangle$. In the contexts of interferometry with light and atoms, e.g., imaging or magnetometry, such states are often known as NOON [109] or GHZ [110] states, respectively. The QFI in this case $\mathbf{Q}_{\psi_{\text{opt}}}(\theta) = (\lambda_+ - \lambda_-)^2$ informs us that the best possible estimate $\hat{\theta}$ for the angle θ will have its variance be lower bounded as

$$\Delta^2 \hat{\theta} \geq \frac{1}{\mathbf{Q}_{\psi_{\text{opt}}}(\theta)} = \frac{1}{(\lambda_+ - \lambda_-)^2}. \tag{2.23}$$

The QFI is additive for repeated measurements and so here and henceforth we consider the QFI *per trial* (i.e., per state probing the parameter of interest in the asymptotic limit; we will always use one probe state per application of the unitary channel). The worst possible scheme, in contrast, uses a probe that remains unchanged by U , such as the pure states $|\pm \mathbf{n}\rangle$ or the maximally mixed state $\mathbb{1}/d$.

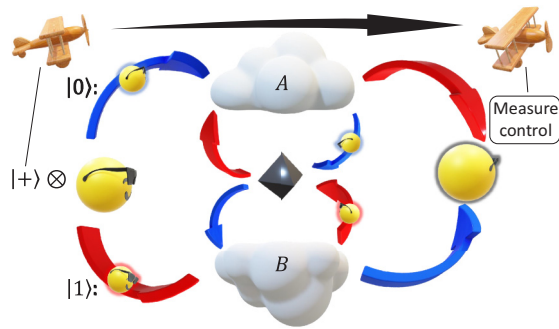


FIG. 1. Schematic for estimating a unitary in the presence of noise. The probe, a maximally mixed state insensitive to unitaries (sunglasses clad), is sent through noise channels (clouds) before and after probing a unitary (diamond). When the control is in state $|0\rangle$ ($|1\rangle$), the probe follows the blue (red) path whose first noise channel is A (B). A superposition-state control leads to indefinite causal order, by way of which a final measurement on the control alone (carried by airplane to circumvent the unitary and noise channels) can learn about the unitary with dramatic advantages over any causally ordered scheme. This scheme can be generalized to higher-dimensional quantum switches that control more than two different orders of operations amongst the three channels here and to more than three channels.

III. ADVANTAGEOUS UNITARY ESTIMATION IN THE PRESENCE OF NOISE USING ICO

In the presence of noise, the QFI tends to decrease, except for some fortuitous situations in which it remains constant. We here consider a general schematic, depicted in Fig. 1, in which some noise affects a probe system both before and after it experiences the unitary transformation, which is a generic scenario where we simply supply different labels for the noise experienced by a probe on either side of a unitary. The order in which the probe traverses the noise and unitary channels can be controlled by a quantum system, again depicted in Fig. 1, such that measuring the control qubit alone allows one to learn about the unitary with a dramatic advantage over any causally ordered scheme. Experimental demonstrations of such quantum control of the order of traversing noise channels have already succeeded [79,80], making the application of this idea to metrology practicable. We here showcase these advantages for three different types of noise: Depolarization, dephasing, and amplitude damping, acting on probes of arbitrarily large dimensions so as to allow for arbitrary unitaries to be estimated.

A. Depolarization noise

We first recapitulate the ICO-driven advantage in estimating the phase of any unitary operation in the presence of strong depolarization noise presented in Ref. [96].

Depolarization noise strongly reduces the ability to estimate θ . Depolarization adds white noise to the state such that, with some probability $1 - p$, one loses all information about the original state and becomes insensitive to all unitary parameters:

$$\mathcal{E}_{\text{depol}}(\rho) = p\rho + (1 - p)\frac{\mathbb{1}}{d}. \quad (3.1)$$

If depolarization occurs either before or after a pure probe state undergoes the unitary transformation, the QFI diminishes as [111]

$$\mathcal{Q}[\mathcal{E}_{\text{depol}}(|\psi\rangle\langle\psi|; \theta)] = \frac{p^2}{p + \frac{1-p}{d/2}} \mathcal{Q}(|\psi\rangle\langle\psi|; \theta), \quad (3.2)$$

where we employ the alternate notation $\mathcal{Q}_\rho(\theta) \leftrightarrow \mathcal{Q}(\rho; \theta)$ when convenient. Convexity of the QFI $\mathcal{Q}(\sum_i p_i \rho_i; \theta) \leq \sum_i p_i \mathcal{Q}(\rho_i; \theta)$ and flatness of the depolarization channel $\mathcal{E}_{\text{depol}}(\sum_i p_i \rho_i) = \sum_i p_i \mathcal{E}_{\text{depol}}(\rho_i)$ lead to the following inequality for all states undergoing depolarization, even including probe states entangled with ancillary quantum systems and joint measurements on the entangled systems:

$$\mathcal{Q}[\mathcal{E}_{\text{depol}}(\rho); \theta] \leq \frac{p^2}{p + \frac{1-p}{d/2}} (\lambda_+ - \lambda_-)^2. \quad (3.3)$$

The resulting minimum variance for any estimate of θ grows as $O(p^{-2})$ whenever the depolarization probability is close to unity (i.e., $1 - p \sim 1$).

It comes as no surprise that depolarizing a probe state both before and after it undergoes a unitary evolution worsens estimates of the unitary's phase. If the two depolarizations have strengths p_A and p_B , the resulting minimum variance grows as $\Delta^2 \hat{\theta} = O(p_A^{-2} p_B^{-2})$. Yet, placing these two depolarizations in a coherently controlled superposition of their causal orders will significantly decrease the estimator variance.

To apply ICO to depolarizing channels, we need a Kraus-operator representation of $\mathcal{E}_{\text{depol}}$. This can be furnished by defining $d^2 + 1$ operators: d^2 two-index operators that provide white noise for a d -element orthonormal basis $\{|n\rangle\}$ by completely mixing up all information,

$$K_{kl}(p) = \sqrt{\frac{1-p}{d}} |k\rangle\langle l|, \quad (3.4)$$

and the identity operator $K_{\mathbb{1}}(p) = \sqrt{p} \mathbb{1}$ that leaves states unchanged.

We use a single application of the unitary channel and two different depolarizing channels, with the orders $\mathcal{E}_{\text{depol}}^{(A)}$ -then- U -then- $\mathcal{E}_{\text{depol}}^{(B)}$ when the control is in state $|0\rangle$ and $\mathcal{E}_{\text{depol}}^{(B)}$ -then- U -then- $\mathcal{E}_{\text{depol}}^{(A)}$ when the control is in state $|1\rangle$. Defining $K_{kl}^{(A)}$ and $K_{mn}^{(B)}$ with p_A and p_B , respectively, we can compute using Eq. (2.14),

$$\begin{aligned} R_{01} &= \sum_{ijkl} \text{Tr}(K_{ij}^{(B)} U K_{kl}^{(A)} \rho K_{ij}^{(B)\dagger} U^\dagger K_{kl}^{(A)\dagger}) \\ &= \frac{p_A(1-p_B)}{d} \text{Tr}(U^\dagger)\langle U \rangle + \frac{p_B(1-p_A)}{d} \text{Tr}(U)\langle U^\dagger \rangle \\ &\quad + \frac{(1-p_A)(1-p_B)}{d^2} + p_A p_B, \end{aligned} \quad (3.5)$$

where expectation values $\langle \cdot \rangle$ are taken with respect to the initial probe state ρ_p . Choosing the least remarkable probe state $\rho_p = \mathbb{1}/d$, which is maximally mixed, possesses the least quantum mechanical properties, and should be insensitive to unitaries because $U \rho_p U^\dagger = \rho_p$ allows one to directly learn

about

$$u = |\text{Tr}(U)|^2 = \left| \sum_i e^{i\lambda_i\theta} \right|^2. \tag{3.6}$$

Since the eigenvalues $\{\lambda_i\}$ of the generators of $SU(N)$ can be readily calculated, this provides a direct window into estimating the unitary's phase θ .

How well can this be done? Defining the $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ basis, starting with the control state in $\rho_c = |+\rangle\langle+|$, then measuring the control state ρ'_c in the $|\pm\rangle$ basis provides an FI equal to the QFI for this state of

$$\begin{aligned} Q_{\text{ICO}}(\theta) &= \frac{(p_A + p_B - 2p_A p_B)^2 \left(\frac{\partial u}{\partial \theta}\right)^2}{d^4 - [(1 - p_A)(1 - p_B) + (p_A + p_B - 2p_A p_B)u + d^2 p_A p_B]^2} \\ &\approx \frac{(p_A + p_B)^2}{d^4 - 1} \left(\frac{\partial u}{\partial \theta}\right)^2 + O(p_A^3, p_B^3, p_A p_B^2, p_A^2 p_B). \end{aligned} \tag{3.7}$$

Because this scales with the second power of the noise and not the fourth, we learn that, for any rotation angle other than $\theta = 0$ or $\theta = \pi$, there always exists a noise threshold above which ICO has an advantage over DCO.

Per the quantum Cramér-Rao bound, the minimum uncertainty on any estimator of θ is given by the inverse of the QFI, showing that this outperforms the best QFI for sensing θ with DCO by a factor on the order of $O(p_A^2, p_B^2, p_A p_B)$:

$$\min_{\text{ICO}} \Delta^2 \hat{\theta} = O\left(\frac{1}{(p_A + p_B)^2}\right) \ll \min_{\text{DCO}} \Delta^2 \hat{\theta} = O\left(\frac{1}{(p_A p_B)^2}\right), \tag{3.8}$$

which provides an essentially unlimited advantage as the depolarization noise increases and p_A and p_B decrease to zero. Even though the expression $\partial u/\partial \theta$ appears in this expression, we have computed the FI in terms of the probability distribution $p_{\pm}(\theta) = \langle \pm | \rho_c | \pm \rangle$ and need not worry about accidentally choosing the optimal estimator for $u(\theta)$ instead of the optimal estimator for θ .

We plot the relative advantage for $d = 2$ in Fig. 2 with $p_A = p_B \equiv p$; when p_A and p_B are different for a given total $p_A + p_B$ or a given fixed $p_A p_B$, the ICO-driven advantage is even greater. These results require only a binary measurement on a single quantum system, as opposed to a generic measurement on a large-dimensional probe state or a joint measurement on the entangled-control-probe state. Similar results can be obtained when the control system begins in any equal-magnitude superposition of $|0\rangle$ and $|1\rangle$ so as to maximize the effect of R_{01} on ρ'_c ; different relative phase choices for the initial control state lead to different optimal measurement bases.

The same result cannot be achieved by entangling the control and probe systems, letting the probe system evolve through one causal sequence, then measuring the final state of the control system. This is because the overall quantum channel that the probe experiences, formed by iterations of the map from Eq. (3.1), preserves the trace of the probe state. As such, tracing out the probe state before or after it evolves does not affect the control state: $\rho_c = \rho'_c$. If either $p_A = 0$ or $p_B = 0$ with DCO, no scheme will be able to estimate U , even with access to ancillary entangled systems. The interference between the different causal sequences is essential (i.e., it is

necessary, but not always sufficient) for probing the properties of the noisy channels.

B. Dephasing noise

Now suppose an alternate noise source, in which the state is subject to dephasing or spin-flip noise. We start by considering a qubit probe state subject to dephasing along a particular axis \mathbf{u} with some probability $1 - p$, characterized by the Pauli matrices $\sigma_{\mathbf{u}} = \mathbf{u} \cdot \boldsymbol{\sigma}$:

$$\mathcal{E}_{\text{dephase}}(\rho) = p\rho + (1 - p)\sigma_{\mathbf{u}}\rho\sigma_{\mathbf{u}}. \tag{3.9}$$

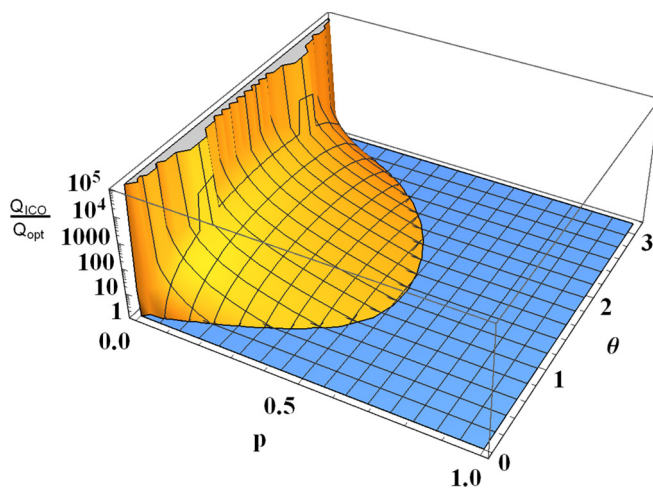


FIG. 2. Advantage of ICO over the best definite-causal-order strategy for estimating a rotation angle of a qubit. Here, Q_{ICO} is given by Eq. (3.7) with $d = 2$, $p_A = p_B \equiv p$, and $u = 4 \cos^2(\theta/2)$; whereas, Q_{opt} is given by the upper bound in Eq. (3.3) with $\lambda_{\pm} = \pm \frac{1}{2}$ and two applications of the depolarizing channel of strength p (i.e., $Q_{\text{opt}} = p^4/[p^2 + 2(1 - p^2)/d]$). Plotted is the increase in QFI versus rotation angle and depolarization noise, which may be interpreted as how many more times a definite-order scheme must be performed to obtain the same precision as an ICO scheme. If we had chosen a different dimension, the dependence on θ would have changed. For θ near $\pi/2$, the advantage persists even with noise level $p > 1/2$, while it grows rapidly and boundlessly with shrinking p . In the blue region of larger p and thus smaller noise, schemes with ICO can be much worse than schemes with definite causal order and so the former should be avoided.

Again considering applications of this channel in a definite order, such as $\mathcal{E}_{\text{dephase}}^{(A)}$ prior to U and $\mathcal{E}_{\text{dephase}}^{(B)}$ afterward, the QFI for a probe state decreases as a function of p_A and p_B . In this situation, the noise channel does not commute with the unitary operation and so the resulting state is not symmetric with respect to $p_A \leftrightarrow p_B$.

Without loss of generality, we fix the axis of the unitary \mathbf{n} to be the z axis of the coordinate system used to define the computational basis for the probe qubit. Then, an optimal probe state without dephasing is $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. With dephasing, the state evolves to a convex mixture of the four linearly dependent pure states $U|\psi\rangle$, $U\sigma_{\mathbf{u}}|\psi\rangle$, $\sigma_{\mathbf{u}}U|\psi\rangle$, and $\sigma_{\mathbf{u}}U\sigma_{\mathbf{u}}|\psi\rangle$. We can calculate the QFI in terms of the eigenvalues ϱ_k and eigenvectors $|\psi_k\rangle$ of ρ through [20]

$$Q_{\rho}(\theta) = \sum_k \frac{1}{\varrho_k} \left(\frac{\partial \varrho_k}{\partial \theta} \right)^2 + 2 \sum_{k \neq l} \frac{(\varrho_k - \varrho_l)^2}{\varrho_k + \varrho_l} \left| \langle \psi_k | \frac{\partial}{\partial \theta} | \psi_l \rangle \right|^2. \tag{3.10}$$

The axis along which the dephasing acts significantly affects the results. With \mathbf{u} along the y axis, for example, the QFI vanishes at $p_A = 1/2$ because the state becomes maximally mixed for all values of θ and p_B . With \mathbf{u} along the x axis, the state becomes independent from p_A , with the QFI remaining independent from both p_A and p_B . As a final example, for \mathbf{u} along the z axis, the QFI vanishes when either $p_A = 1/2$ or $p_B = 1/2$, taking the form $(\frac{1}{2} - p_A)^2 (\frac{1}{2} - p_B)^2$. Some of these can be dramatically outperformed by schemes with ICO and some cannot.

To introduce ICO, we need only the Kraus operators $K_{\mathbb{1}}^{(O)} = \sqrt{p_O} \mathbb{1}$ and $K_{\mathbf{u}}^{(O)} = \sqrt{1 - p_O} \sigma_{\mathbf{u}}$. As before, we calculate using Eq. (2.14),

$$\begin{aligned} R_{01} &= \sum_{ij} \text{Tr} (K_i^{(B)} U K_j^{(A)} \rho K_i^{(B)\dagger} U^\dagger K_j^{(A)\dagger}) \\ &= p_A(1 - p_B)s + p_B(1 - p_A)s^* \\ &\quad + (1 - p_A)(1 - p_B) + p_A p_B, \end{aligned} \tag{3.11}$$

where we have defined

$$s(\theta, \mathbf{u}; \rho_p) = \langle \sigma_{\mathbf{u}} U^\dagger \sigma_{\mathbf{u}} U \rangle. \tag{3.12}$$

How does the amount of information about θ in ρ'_c compare to the QFI for probe states with DCO? Whereas, the QFI often vanishes for probe states with DCO when $p_A = p_B = 1/2$, R_{01} and therefore ρ'_c retains information about θ at $p_A = p_B = 1/2$, so long as s depends on θ . One can compute $\sigma_{\mathbf{u}} U^\dagger \sigma_{\mathbf{u}} U$ to find it independent from the azimuthal angle of \mathbf{u} . The quantity s depends on θ so long as $|u_z| \neq 1$; thus, for any probe state ρ_p , even a maximally mixed probe state, one can learn about θ using ICO. When \mathbf{u} was along the y axis and $p_A = p_B = 1/2$, we saw that the QFI was zero for DCO; whereas, here it is

$$Q_{\text{ICO}}(\theta) = \frac{1}{3 - 2 \text{Re}(s) - \text{Re}^2(s)} \left[\frac{\partial \text{Re}(s)}{\partial \theta} \right]^2, \tag{3.13}$$

with $s = \langle U^{\dagger 2} \rangle$. This constitutes an *infinite advantage* (an infinite increase in the QFI ratio) in this particular scenario: Even a maximally mixed probe state, with $s = \cos \theta$, can provide

nonzero information about θ in a situation in which DCO provides zero information about θ .

The treatment here can be repeated for arbitrary dimensions by replacing $\sigma_{\mathbf{u}}$ by some other unitary operation. For some channel,

$$\mathcal{E}_{\text{dephase}}(\rho) = p\rho + (1 - p)V\rho V^\dagger, \tag{3.14}$$

the final result for R_{01} remains the same, now with

$$s = \langle V^\dagger U^\dagger V U \rangle. \tag{3.15}$$

So long as U and V do not commute, this provides information about θ to the control state ρ'_c that can be measured. In fact, the term $s = \langle V^\dagger U^\dagger V U \rangle$ is also connected to out-of-time-ordered correlators that characterize quantum information scrambling and, through it, the celebrated Kirkwood-Dirac distribution [112] (see Ref. [113] for further applications of the quantum switch for measuring noncommutativity).

Whether or not there is an *advantage* from ICO depends on whether or not schemes with DCO lose all information from such dephasing. Considering a spin system, with U performing an $SU(2)$ rotation of a spin- J particle, all previous calculations hold true with $\theta \mapsto 2J\theta$ for generalized dephasing V enacting a π rotation about some axis \mathbf{u} . This means that one can again attain an infinite advantage in estimating θ using ICO for dephasing along the y axis, even using a maximally mixed probe state, relative to the optimal quantum strategy of using a pure superposition of extremal eigenstates of U (i.e., NOON- or GHZ-type states of the correct orientation).

Next, considering more general dephasing operators V , we can speculate on a large class of ICO-driven advantages. In large dimensions and for all but pathological cases of U and V , the four pure states $U|\psi\rangle$, $UV|\psi\rangle$, $VU|\psi\rangle$, and $VUV|\psi\rangle$ are linearly independent, even though they were dependent for the qubit case. If they are all orthogonal, the four eigenvalues of ρ after evolving through the unitary and pair of dephasing channels are $p_A p_B$, $p_A(1 - p_B)$, $p_B(1 - p_A)$, and $(1 - p_A)(1 - p_B)$; these cause the QFI to identically vanish at $p_A = p_B = 1$ in Eq. (3.10). Since the ICO-evolved state ρ'_c continues to depend on θ through s , these could present another array of ICO-driven advantages for estimation of unitaries in the presence of noise.

C. Amplitude damping noise

Now consider a final type of noise, in which the probe has a propensity to relax from some excited state to its ground state. This well-known amplitude damping channel [114] acting on a qubit

$$\mathcal{E}_{\text{amp. damp.}} \left[\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \right] = \left[\begin{pmatrix} \rho_{00} + (1 - p)\rho_{11} & \rho_{01}\sqrt{p} \\ \sqrt{p}\rho_{10} & p\rho_{11} \end{pmatrix} \right] \tag{3.16}$$

is characterized by the two Kraus operators

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{p} \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & \sqrt{1 - p} \\ 0 & 0 \end{pmatrix}, \tag{3.17}$$

that cause a system to relax toward state $|0\rangle$ when p gets closer to 0. Again considering a unitary along the z axis and the optimal probe state $|+\rangle$, the QFI for measuring the

unitary’s phase with amplitude damping both before and after application of the unitary degrades to

$$\mathbf{Q}_{|+\rangle}(\theta) = p_A p_B. \quad (3.18)$$

Considering the general case of fixed causal order where the initial probe state is arbitrary, the evolved state is

$$\mathcal{E}(\rho) = \begin{pmatrix} 1 - p_A p_B (1 - \rho_{00}) & e^{i\theta} \sqrt{p_A p_B} \rho_{01} \\ e^{-i\theta} \sqrt{p_A p_B} \rho_{10} & p_A p_B (1 - \rho_{00}) \end{pmatrix}. \quad (3.19)$$

This has a QFI of $\mathbf{Q} = 4p_A p_B |\rho_{01}|^2$, so maximal information about θ is obtained when $|\rho_{01}|$ is maximized, confirming our intuition that $|+\rangle$ remains an optimal probe state in the presence of noise.

With a quantum switch controlling the orders of applications of two amplitude damping channels on a maximally mixed state, we can readily compute

$$R_{01} = \frac{1}{2} [1 - e^{i\theta} (p_A - 1) \sqrt{p_B} - e^{-i\theta} \sqrt{p_A} (p_B - 1) + p_A p_B]. \quad (3.20)$$

The QFI for the evolved control state is a bit involved, though it is simply given by Eq. (3.10) and the eigensystem of a 2×2 matrix, so we write the results in the relevant limit of neglecting terms of order $O(p_A^{3/2}, p_B^{3/2}, p_A p_B^{1/2}, p_A^{1/2} p_B)$:

$$\mathbf{Q}_{\text{ICO}}(\theta) \approx \frac{(\sqrt{p_A} - \sqrt{p_B})^2}{4} + \frac{\sin^2 \theta}{12} (p_A + p_B + 14\sqrt{p_A p_B}). \quad (3.21)$$

This again provides an unlimited benefit in terms of QFI ratio or minimum uncertainty ratio relative to all schemes with DCO in the limit of small p_A and p_B and a formally infinite advantage when either p_A or p_B vanishes (this can also be seen because schemes with DCO leave the probe state independent from θ when p_A or p_B vanishes). Notably, this advantage can be attained for any unitary, even when $\theta = 0$, with the exception of $p_A = p_B$ when $\theta = 0$.

Another type of amplitude damping channel has $\mathcal{E}^{(A)}$ sending a system toward state $|0\rangle$ and $\mathcal{E}^{(B)}$ toward $|1\rangle$; different relaxation tendencies occur on different sides of the unitary. Mathematically, this happens when $\mathcal{E}^{(A)}$ has the Kraus operators from before and $\mathcal{E}^{(B)}$ has Kraus operators

$$K_0^{(B)} = \begin{pmatrix} \sqrt{p} & 0 \\ 0 & 1 \end{pmatrix}, \quad K_1^{(B)} = \begin{pmatrix} 0 & 0 \\ \sqrt{1-p} & 0 \end{pmatrix}. \quad (3.22)$$

In this case, the QFI for DCO schemes with unitary U about the z axis and optimal probe state $|+\rangle$ again takes the form $\mathbf{Q}_{|+\rangle}(\theta) = p_A p_B$ and is optimal among definite-order schemes, as can again be recognized from the DCO QFI $\mathbf{Q} = 4p_A p_B |\rho_{01}|^2$. In contrast, the small- p limit of the QFI for schemes with ICO is again $\mathbf{Q}_{\text{ICO}}(\theta) \approx (\sqrt{p_A} - \sqrt{p_B})^2/4$.

Amplitude damping toward some state $|0\rangle$ can be extended to amplitude damping occurring identically on $n = \log_2 d$ qubits in parallel. In the case of ICO, R_{01} simply gets modified as $R_{01} \mapsto R_{01}^n$, retaining the dependence on $\sqrt{p_A}$ and $\sqrt{p_B}$ to first order as

$$R_{01} \approx \frac{1 + e^{i\theta} n \sqrt{p_B} + e^{-i\theta} n \sqrt{p_A}}{2^n}. \quad (3.23)$$

The QFI then becomes, to lowest order in p_A and p_B ,

$$\mathbf{Q}_{\text{ICO}}(\theta) = \frac{n^2 \sin^2 \theta}{4^n - 1} (\sqrt{p_A} + \sqrt{p_B})^2 + \frac{n^2 \cos^2 \theta}{4^n} (\sqrt{p_A} - \sqrt{p_B})^2. \quad (3.24)$$

This worsens with n but retains the same scaling with p_A and p_B as for our depolarization case in Sec. III A. The optimal definite-order scheme now involves entangled qubits in a GHZ state $(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$, but amplitude damping diminishes its QFI by $p_A^n p_B^n$:

$$\mathbf{Q}_{\text{GHZ}}(\theta) = 2 \frac{p_A^n p_B^n}{1 + (1 - p_A p_B)^n + p_A^n p_B^n}. \quad (3.25)$$

A better definite-order scheme might be to use n qubits in parallel, each in the $|+\rangle$ state, which would allow the QFI to simply scale with n instead of diminishing exponentially. Still, such definite-order schemes are limited to $\mathbf{Q} \sim O(p_A p_B)$, while ICO allows $\mathbf{Q}_{\text{ICO}} \sim O(p_A, p_B, \sqrt{p_A p_B})$ for any number of qubits n [115]. We again see a general advantage for ICO over definite-order schemes for estimation of a unitary in the presence of noise, even using maximally mixed states as inputs, with the advantage growing with the amount of noise and diminishing with the dimension of the probe system.

IV. MULTIPARAMETER ESTIMATION

One notices above that the control state depends not only on the parameters of the unitary being estimated but also on the strength of the noise channels. As such, one can imagine using the same protocol to estimate the noise levels instead of the unitary’s parameters. One cannot estimate both simultaneously because the control state only depends on a single parameter, R_{01} , through which both θ and p are to be estimated. It then follows that higher dimensional control states that depend on more parameters may be used to simultaneously estimate multiple parameters of the quantum channels, as we presently show.

A. Estimation of depolarization noise and unitary phase

We now step into the world of multiparameter estimation. For a measurement only of the control to yield information about more than one parameter, it must have more than one functional dependence on those parameters. In the examples above, we only had access to the parameter R_{01} , which was often real, notably in the case of depolarization channels with maximally mixed probe states. Here we show how using control systems with larger dimensions for the quantum switch allows one to simultaneously estimate both depolarization noise strength and the unitary channel’s phase.

Referring again to Fig. 1, there are three total channels: Two depolarizations and one unitary. The six different orders of traversing these channels can be combined to give different functional dependencies on p_A , p_B , and $u(\theta)$ such that the noise and unitary parameters can be simultaneously estimated. We will not require all six orders to determine only three parameters, so we choose the three orders $\mathcal{E}^{(A)} \circ \mathcal{E}^{(B)} \circ U$, $U \circ \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}$, and $\mathcal{E}^{(B)} \circ U \circ \mathcal{E}^{(A)}$ when the control is in state $|0\rangle$, $|1\rangle$, and $|2\rangle$, respectively. We then calculate using

Eqs. (2.12) and (2.14) for the three orders:

$$\begin{aligned} R_{01} &= p_A p_B + \frac{1 - p_A p_B}{d} \langle U \rangle \text{Tr}(U^\dagger), \\ R_{02} &= p_A + \frac{(1 - p_A)(1 - p_B)}{d^2} + \frac{p_B(1 - p_A)}{d} \langle U \rangle \text{Tr}(U^\dagger), \\ R_{12} &= p_B + \frac{(1 - p_A)(1 - p_B)}{d^2} + \frac{p_A(1 - p_B)}{d} [\langle U \rangle \text{Tr}(U^\dagger)]^*. \end{aligned} \quad (4.1)$$

These three different functional dependencies on p_A , p_B , and $\langle U \rangle \text{Tr}(U^\dagger)$ allow all three to simultaneously be estimated from the evolved control state ρ'_c , even though there was only one copy of each channel being probed. This again holds even if the probe state is maximally mixed and therefore insensitive to each of p_A , p_B , and U for definite-order schemes. When the probe is maximally mixed, we again find the dependence on θ for the ICO scheme through $u(\theta) = |\text{Tr}(U)|^2 = d \langle U \rangle \text{Tr}(U^\dagger)$. Moreover, dependence on θ is maintained even when $p_A = p_B = 0$, showing that the ability to completely control the order of the depolarization and unitary channels (not restricted to the unitary always occurring between the two depolarizations) leads to sensitivity that is even more impossible with DCO.

Measuring the control state in the $(|i\rangle \pm |j\rangle)/\sqrt{2}$ basis directly yields the real part of R_{ij} , where R_{ij} is automatically real when the probe is maximally mixed. This can be facilitated by a POVM with six elements $(|i\rangle \pm |j\rangle)(\langle i| \pm \langle j|)/4$, $i < j$, where the extra factor of 2 is required for normalization. The six probabilities are

$$P_{ij\pm} = \frac{1 \pm R_{ij}}{6}, \quad i < j, \quad (4.2)$$

and we can use them to calculate the QFI matrix. Incredibly, this matrix is nonzero and invertible even when $p_A = p_B = 0$, where *no* definite-order strategy could ever determine u . We write the FI matrix in the $\theta = (\theta, p_A, p_B)$ basis and record the result for $p_A = p_B = 0$, with the full expression given in the Appendix A:

$$\mathbf{F}_{\text{ICO}}(\theta) = \begin{pmatrix} \frac{(\partial u / \partial \theta)^2}{3d^4 - 3u^2} & 0 & 0 \\ 0 & \frac{d^4 - 2d^2 + (u-2)u + 2}{3(d^4 - 1)} & \frac{2(u-1)}{3(d^2 + 1)} \\ 0 & \frac{2(u-1)}{3(d^2 + 1)} & \frac{d^4 - 2d^2 + (u-2)u + 2}{3(d^4 - 1)} \end{pmatrix}. \quad (4.3)$$

We can call this a *formally infinite* advantage due to the unlimited increase in FI in simultaneously estimating three parameters using ICO, even using a maximally mixed probe for the latter, valid for unitaries and depolarization channels in arbitrary dimensions d . Of course, there might exist another measurement strategy that coaxes even more information from ρ'_c , as we have not computed the QFI matrix for this state, but we are satisfied with an infinite increase in QFI relative to any strategy with DCO, especially because our FI matrix is attainable through a straightforward measurement procedure. Moreover, since we have chosen a fixed POVM to obtain these results, we need not worry about factors of f from Eq. (2.20) and are guaranteed to saturate the classical Cramér-Rao bound for the minimum uncertainties of each parameter [Eq. (2.15)] in the asymptotic limit.

We can also inspect the large- d limit, which makes it more difficult to estimate θ as seen before. Even in this limit, each of p_A and p_B can be estimated without much trouble, given the constant term in $[d^4 - 2d^2 + (u-2)u + 2]/[3(d^4 - 1)] = (1 - 2/d^2)/3 + O(d^{-4})$. This prompts us to calculate the $d \rightarrow \infty$ limit of \mathbf{F} for arbitrary p_A and p_B , perhaps having in mind a measurement with macroscopic probe systems. Different functional forms $u(\theta)$ will behave differently in the limit of large d , so we inspect only the (p_A, p_B) submatrix of \mathbf{F} to show how it allows p_A and p_B to simultaneously be estimated:

$$\lim_{d \rightarrow \infty} \mathbf{F}_{\text{ICO}}(p_A, p_B) = \begin{pmatrix} \frac{(1-2p_A^2)p_B^2+1}{3(p_A^2-1)(p_A^2p_B^2-1)} & \frac{p_A p_B}{3-3p_A^2 p_B^2} \\ \frac{p_A p_B}{3-3p_A^2 p_B^2} & \frac{(1-2p_B^2)p_A^2+1}{3(p_B^2-1)(p_A^2 p_B^2-1)} \end{pmatrix}. \quad (4.4)$$

Just as for single-parameter estimation, probe states other than the maximally mixed state can be used to investigate other properties of U . There is still only one functional dependence on the unitary's parameters through $\langle U \rangle \text{Tr}(U^\dagger)$, so this parameter could simultaneously be estimated alongside p_A and p_B if one desires. The real part of $\langle U \rangle \text{Tr}(U^\dagger)$ would be accessible through the POVM described in this section, while a more general POVM might gain access to the imaginary part at the same time.

To conclude this section, we note that one could have chosen other combinations of orders to estimate these three parameters. We tabulate in the Appendix B all of the matrix elements R_{ij} that would arise from all 36 interference terms of the six possible orders of traversing the two depolarization channels and one unitary channel. One could use higher-dimensional control states to gain redundant information about the parameters of interest because these 36 elements have more than three functional dependencies on p_A , p_B , and $u(\theta)$. These interference terms are responsible for ICO's advantages in metrology.

B. Estimating dephasing noise and unitary phase

Consider the same three orders as above but replace the depolarizing channels with dephasing channels in the direction \mathbf{u} pointing anywhere along the equator for qubit rotations about the z axis. The three off-diagonal elements of ρ'_c can be calculated using

$$\begin{aligned} R_{01} &= (p_A + p_B - 2p_A p_B) \cos \theta + 2p_A p_B - p_A - p_B + 1, \\ R_{02} &= p_A(1 - \cos \theta) + \cos \theta, \\ R_{12} &= p_B(1 - \cos \theta) + \cos \theta. \end{aligned} \quad (4.5)$$

These three linearly independent functions then facilitate the simultaneous estimation of θ and the noise strengths p_A and p_B using the six-element POVM with projections of the control state onto the $(|i\rangle \pm |j\rangle)/\sqrt{2}$ basis.

The FI matrix is again a complicated expression, so we here record the result for $p_A = p_B = 1/2$, where definite-order strategies cannot be used to estimate θ . In this limit,

even the inverse is not too unwieldy, becoming

$$\mathbf{F}_{\text{ICO}}^{-1}(\theta, p_A, p_B) = \begin{pmatrix} \frac{\cos\theta+1}{\cos\theta+3} & -\frac{2\sin\theta}{3\cos\theta+9} & -\frac{2\sin\theta}{3\cos\theta+9} \\ -\frac{2\sin\theta}{3\cos\theta+9} & \frac{4-4\cos\theta}{3\cos\theta+9} & 0 \\ -\frac{2\sin\theta}{3\cos\theta+9} & 0 & \frac{4-4\cos\theta}{3\cos\theta+9} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 3 + \frac{6}{\cos\theta+1} & \frac{3}{2}(\cos\theta+3)\csc\theta & \frac{3}{2}(\cos\theta+3)\csc\theta \\ \frac{3}{2}(\cos\theta+3)\csc\theta & 3\csc^2\frac{\theta}{2} - \frac{3}{2} & \frac{3}{8}(\cos\theta+3)\csc^2\frac{\theta}{2} \\ \frac{3}{2}(\cos\theta+3)\csc\theta & \frac{3}{8}(\cos\theta+3)\csc^2\frac{\theta}{2} & 3\csc^2\frac{\theta}{2} - \frac{3}{2} \end{pmatrix}. \tag{4.6}$$

We place the full expression for the FI matrix in the Supplemental Material, with a determinant that only vanishes when $(1 - p_A)(1 - p_B) \sin(\theta/2) \sin\theta = 0$. As with depolarization, the bound of Eq. (2.15) can be saturated without resorting to considerations of QFI from Eq. (2.20) because we have chosen a fixed, accessible measurement scheme. We can again conclude that ICO may grant a formally infinite advantage over definite order schemes in this multiparameter estimation context.

C. Estimating amplitude damping noise and unitary phase

Next let us show how to simultaneously estimate both amplitude damping channels' noise parameters and the unitary's phase by controlling the order of operations with a higher-dimensional quantum switch. Keeping the same nominal three orders as in the previous sections, we use the maximally mixed qubit probe state to calculate

$$\begin{aligned} R_{01} &= \frac{1}{2}[e^{-i\theta}(1 - p_A p_B) + p_A p_B + 1], \\ R_{02} &= \frac{1}{2}[-(p_A - 1)\sqrt{p_B}e^{-i\theta} + p_A p_B - \sqrt{p_A}(p_B - 1) + 1], \\ R_{12} &= \frac{1}{2}[-\sqrt{p_A}(p_B - 1)e^{i\theta} + p_A p_B - (p_A - 1)\sqrt{p_B} + 1]. \end{aligned} \tag{4.7}$$

We can again measure the evolved control state using the POVM comprised of projectors onto states $(|i\rangle \pm |j\rangle)/\sqrt{2}$ with $i < j$. This will be sensitive to the real parts of R_{ij} , which are all that we require for estimating θ , p_A , and p_B . Computing the FI matrix with this method yields complicated expressions, so we plot the appropriate component of the inverse $(\mathbf{F}^{-1})_{\theta\theta}$ in Fig. 3 to display the phase sensitivity of the multiparameter scheme as a function of θ and $p_A = p_B \equiv p$. The minimum variance $\Delta^2\hat{\theta}$ is seen to be bounded, showing that this multiparameter estimation scheme is successful, and significantly outperforms the limit one could achieve with single-parameter definite-order schemes $[1/p^2; \text{cf. Eq. (3.18)}]$ when p and θ are small. We have license to use the components of \mathbf{F}^{-1} as the covariances on the estimated parameters due to having supplied a fixed POVM that can saturate the bound of Eq. (2.15). This comparison is the same whether we use a multiparameter or a single-parameter estimation scheme for the qubit probe with DCO because the off-diagonal components $\mathbf{Q}_{\theta p_A}$ and $\mathbf{Q}_{\theta p_B}$ vanish in that case and so multiparameter estimation does not worsen the estimation.

D. Estimation of a unitary's phase and axis

What if one desires to simultaneously estimate more than one parameter from U using ICO? Increasing the dimension of the control system will again be necessary, but we saw above that depolarization channels with maximally mixed probes only give access to one parameter from U (Appendix A shows all of the dependencies to be $u(\theta) = |\text{Tr}(U)|^2$ and $\tilde{u}(\theta, \mathbf{n}; \rho) = \langle U \rangle \text{Tr}(U^\dagger)$, so one could consider engineering more complicated probe states to glean information about an additional function of (θ, \mathbf{n}) through \tilde{u} ; the experimental challenge of creating other probe states must be balanced with their effectiveness at identifying requisite parameters and we leave such study to further work). Here we explore whether ICO in the presence of dephasing or amplitude damping channels, which do not act isotropically on a state, gives access to more parameters of U for simultaneous estimation. It turns out that the former is insufficient while the latter can be used for such simultaneous estimation with maximally mixed probe states.

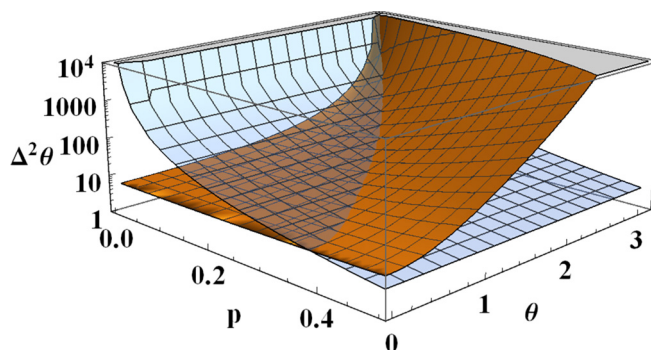


FIG. 3. Minimum variance that saturates the Cramér-Rao bound for estimating a phase θ of a qubit rotation subject to amplitude damping noise for our ICO (orange, solid) and the optimal definite-order (blue, translucent) multiparameter estimation schemes. The multiparameter schemes requires inverting the full Fisher information matrix and inspecting the $\theta\theta$ component thereof, which is plotted as $\Delta^2\theta$ here. Our scheme with ICO is significantly better than strategies with definite causal order when p and θ are small and vice versa; ICO can be used to great effect in the former regime and should be avoided in the latter. These uncertainties are plotted for a single trial and will be normalized by the total number of independent trials, so one need only worry about the ratio between the variances for ICO and definite causal order, with less concern paid to the absolute magnitudes.

1. Dephasing channels cannot be used to estimate phase and axis

Consider the case of qubit rotations, in which we parametrize the unitary's axis as $\mathbf{n} = (\sin \Theta \cos \Phi, \sin \Theta \sin \Phi, \cos \Theta)$. Subjecting a maximally mixed probe state to dephasing noise as above with a coherent control of the three orders of the channels yields the following matrix elements when we consider dephasing along the z axis [i.e., $\mathcal{E}(\rho) = p\rho + (1-p)\sigma_z\rho\sigma_z$]:

$$\begin{aligned} R_{01} &= \frac{1}{2}[(\cos 2\Theta + 2 \sin^2 \Theta \cos \theta)(-2p_A p_B + p_A + p_B) \\ &\quad + 2p_A p_B - p_A - p_B + 2], \\ R_{02} &= \frac{1}{2}[-(p_A - 1)(\cos 2\Theta + 2 \sin^2 \Theta \cos \theta) + p_A + 1], \\ R_{12} &= \frac{1}{2}[-(p_B - 1)(\cos 2\Theta + 2 \sin^2 \Theta \cos \theta) + p_B + 1], \\ R_{24} &= \frac{1}{2}\{(\cos 2\Theta + 2 \sin^2 \Theta \cos \theta)[p_A(p_B - 1) \\ &\quad + \sqrt{(p_A - 1)p_A(p_B - 1)p_B} - p_B + 1] + 3p_A p_B \\ &\quad + 3\sqrt{(p_A - 1)p_A(p_B - 1)p_B} - p_A - p_B + 1\}. \end{aligned} \quad (4.8)$$

We have included an additional ordering by adding a control state $|4\rangle$ that sends the probe through the channels as $\mathcal{E}^{(B)} \circ \mathcal{E}^{(A)} \circ U$ to showcase a general trend (keeping the same

$$\begin{aligned} R_{01} &= \frac{1}{2}[\sin^2 \Theta \sqrt{p_A p_B} + p_A p_B \cos^2 \Theta - \cos \theta (\sin^2 \Theta \sqrt{p_A p_B} + p_A p_B \cos^2 \Theta - 1) + i \cos \Theta (p_A p_B - 1) \sin \theta + 1], \\ R_{02} &= \frac{1}{8}\{2\sqrt{p_B}[\cos 2\Theta (p_A \sqrt{p_B} - 2\sqrt{p_A p_B} + p_A + \sqrt{p_B} - 1) \sin^2(\theta/2) + 2i(p_A - 1) \cos \Theta \sin \theta] \\ &\quad + [(\sqrt{p_A} - 1)^2 p_B - 3(p_A - 1)\sqrt{p_B}] \cos \theta - 2\sqrt{p_A}(p_B - 2) + p_A(3p_B - \sqrt{p_B}) - p_B + \sqrt{p_B} + 4\}, \\ R_{12} &= \frac{1}{8}\{2 \cos 2\Theta [2p_A(p_B - \sqrt{p_B}) + \sqrt{p_A}(p_B - 1) - p_B + 1] \sin^2(\theta/2) - 4i\sqrt{p_A}(p_B - 1) \cos \Theta \sin \theta \\ &\quad + [-3\sqrt{p_A}(p_B - 1) + 2p_A(p_B - \sqrt{p_B}) - p_B + 1] \cos \theta + (2p_A - \sqrt{p_A} + 1)p_B - 2(p_A - 2)\sqrt{p_B} + \sqrt{p_A} + 3\}. \end{aligned} \quad (4.9)$$

From these expressions, we see the importance of the interplay between the particular amplitude damping channel and \mathbf{n} : Φ is absent from R_{ij} . ICO with this particular amplitude damping channel can be used to simultaneously estimate the unitary's phase and the polar angle of its rotation axis, while another amplitude damping channel that singles out a different preferred axis could be used to learn about another projection of \mathbf{n} .

Suppose one wishes to simultaneously estimate both noise parameters p_A and p_B in addition to the two unitary parameters θ and Θ using ICO and this pair of amplitude damping channels. One must immediately be wary, as we have only computed three quantities R_{ij} and seek four parameters. There are a few paths forward: (a) one can perform a measurement with different POVM elements sensitive to the real and imaginary parts of R_{ij} , using projections onto the states $(|i\rangle \pm i|j\rangle)/\sqrt{2}$ in addition to $(|i\rangle \pm |j\rangle)/\sqrt{2}$; (b) one can consider situations in which the two noise levels are known to be equal, $p_A = p_B \equiv p$, such that the total number of parameters to be estimated is three; (c) one may seek to only estimate a subset of the parameters, implicitly assuming the rest to be known; or, (d) one can consider expanding the dimension of the control system, such as by adding a control state $|4\rangle$ that sends the probe through the channels as $\mathcal{E}^{(B)} \circ \mathcal{E}^{(A)} \circ U$

orders as in Appendix C). Again, these are independent from Φ due to the particular dephasing axis; another dephasing axis allows one to inspect other projections of \mathbf{n} onto that axis. The only angular information, however, arises in the form of the single function $\sin^2 \Theta \cos \theta + \cos 2\Theta$. This function is indeed sensitive to the rotation angle (phase) and the projection of the dephasing axis onto the unitary's rotation axis, with this projection explaining why ICO could be used above for x - and y -axis dephasings but not z -axis dephasing for unitaries about the z axis. Since there is only one function present, only one variable can be estimated. If the unitary's rotation angle is known, this can be used to estimate the rotation axis and vice versa, but under no circumstances can this be used to estimate two unitary parameters simultaneously. Probe states that are not maximally mixed would be necessary to perform such a simultaneous estimation with ICO.

2. Amplitude damping channel can be used to estimate phase and axis

Consider again the case of qubit rotations with a general axis \mathbf{n} . Subjecting a maximally mixed probe state to amplitude damping noise as above with a coherent control of the three orders of the channels yields the matrix elements

(keeping the same orders as in the Appendix A), which provides new functions of the four parameters such as

$$R_{04} = \frac{1}{2}(p_A(p_B - \sqrt{p_B}) - \sqrt{p_A}(p_B - 1) + \sqrt{p_B} + 1). \quad (4.10)$$

We now inspect the performance of measuring the control state in the $(|i\rangle \pm |j\rangle)/\sqrt{2}$ basis as before. We consider the case where $p_A = p_B \equiv p$ to streamline the assessment, using only the coefficients from Eq. (4.9). Normalizing the minimum values of $\Delta^2 \theta$ and $\Delta^2 \Theta$ by the inverse of Eq. (3.18), which is the increase in uncertainty one would expect for strategies with DCO, we plot the minimum uncertainties for θ , Θ , and p in Figs. 4–6, respectively for various small values of p . As above, we consider the components of the inverse of the FI matrices to represent the minimum uncertainties, which is justified in the asymptotic limit of saturating Eq. (2.15) with a fixed POVM. As discussed in the figure captions, one can observe a significant advantage relative to definite-order schemes for estimating θ and Θ when p is small and the former two parameters are in the proper regimes, with the advantage qualitatively corresponding to $O(p^2)$ smaller variances, while one can sensibly estimate p at the same time if θ is small.

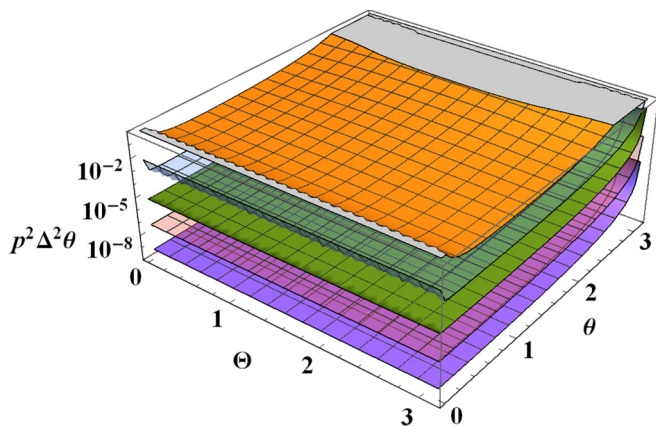


FIG. 4. Decrease in uncertainty for estimating the rotation angle θ of a qubit rotation when simultaneously estimating the unitary’s rotation angle, its axis’s polar coordinate, and an amplitude damping noise level p using ICO, relative to schemes with definite causal order that achieve a minimum p^2 . This is equal to the ratio of the smallest possible inverse of the FI matrix without ICO, p^2 , to the θ, θ element of the inverse of the measured FI matrix with ICO. The probe, which for ICO is a maximally mixed state, goes through an amplitude damping noise channel with strength p both before and after the unitary. The different sheets plotted correspond to p ranging from 10^{-1} to 10^{-5} by factors of 10, with the p increasing from the lowest to the highest sheet; the advantage is approximately $O(p^2)$ smaller variances. The upper cutoff is set to 1 to single out the regime of ICO-driven advantages. The advantage is most prominent when θ is further from 0 and π .

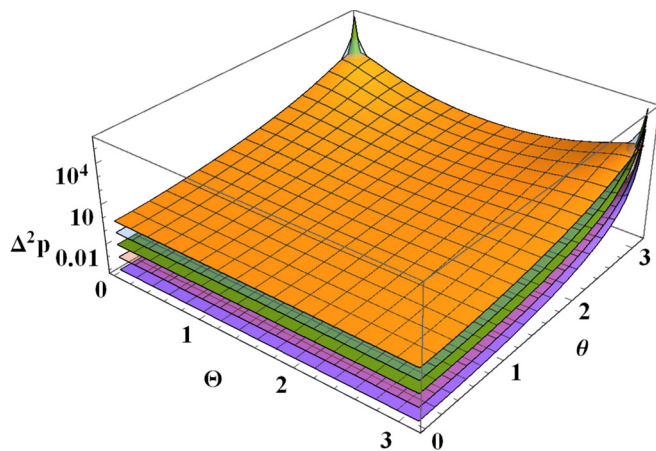


FIG. 6. Same as Figs. 4 and 5 but the uncertainty is plotted for estimating the noise level of the amplitude damping channel and is not normalized. Again, p increases from the lowest to the highest sheet. The uncertainty is lowest when θ is smallest.

V. MULTIPARAMETER ESTIMATION WITH MULTIPLE COPIES OF IDENTICAL NOISY UNITARIES

Our above analyses used ICO to crucially control the order in which a unitary and noise channels were applied, schematized in Fig. 1. Other studies of ICO for noisy metrology, in contrast, assumed multiple identical copies of the same noisy channel, without the possibility of controlling the order of the noise and unitary within one joint channel. An example

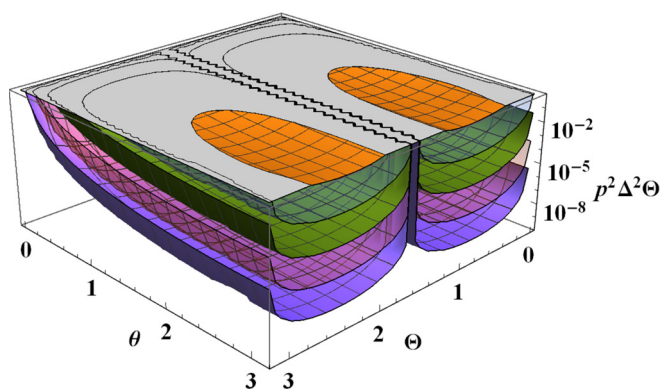


FIG. 5. Same as Fig. 4, but the uncertainty is plotted for estimating the polar angle Θ of a qubit rotation’s rotation axis. Again, p increases from the lowest to the highest sheet with approximate advantages for ICO of the order $O(p^2)$. Now the advantage is most prominent when θ is further from 0 when Θ is furthest from 0, $\pi/2$, and π .

scheme can be seen in Fig. 7, where now each one unitary is embedded in noise channel A and another identical unitary in noise channel B , with ICO merely controlling the order of overall channels A and B . For such schemes, no information about the unitary can be found if the noise channels are completely depolarizing or completely amplitude damping, in contrast to our earlier schemes, even in the limit of large numbers of copies of the channels [72,116]. In this section, we show how identical-channel schemes with fixed causal orders *within* each channel can be extended to multiparameter estimation in arbitrary dimensions using ICO. We also show how such strategies can retain FI of order $O(p)$ for any number D depolarization channels, even though naive schemes with DCO would have FI dramatically lower at order $O(p^D)$; even though such an advantage should also be attainable in the limit of arbitrary copies of the channels by using adaptive techniques or ancilla-entangled strategies [116], we provide an explicit procedure to attain such an advantage here.

Consider the joint unitary-depolarization channel

$$\mathcal{E}_{U\text{-depol}}(\rho) = pU\rho U^\dagger + (1-p)\frac{\mathbb{1}}{d}. \tag{5.1}$$

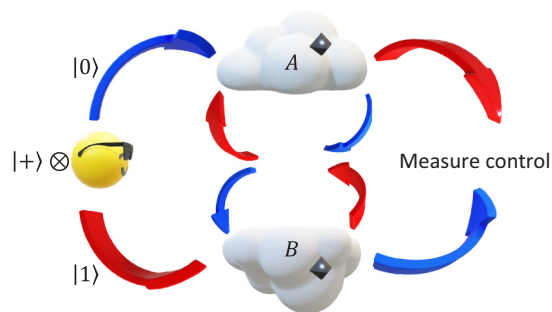


FIG. 7. Schematic for metrology with ICO, given two copies A and B of a noisy unitary channel. As in Fig. 1, a control system dictates the order in which the channels are traversed by the probe, while a measurement on the control alone that has not interacted with the noisy channels is sufficient to infer properties of the channels.

This can be achieved by concatenating the unitary and depolarization channels above in either fixed order (unitary then depolarization or depolarization then unitary), so its Kraus operators can be chosen to be

$$K_{kl}(p) = \sqrt{\frac{1-p}{d}} |k\rangle\langle l|U \tag{5.2}$$

and $K_{\mathbb{1}} = \sqrt{p}U$.

$$\begin{aligned} R_{01} &= p_{APBPC} + \frac{p_{APB} + p_{APC} + p_{BPC} - 3p_{APBPC}}{d} \text{Tr}(U^2)\langle U^{\dagger 2} \rangle + \frac{1 - p_{APB} - p_{APC} - p_{BPC} + 2p_{APBPC}}{d^2} \\ &= p^3 + 3p^2 \frac{1-p}{d} \text{Tr}(U^2)\langle U^{\dagger 2} \rangle + \frac{1 - 3p^2 + 2p^3}{d^2}, \\ R_{02} &= p_{APBPC} + \frac{p_A(1 - p_{BPC})}{d} \text{Tr}(U)\langle U^{\dagger} \rangle + \frac{(1 - p_A)p_{BPC}}{d} \text{Tr}(U^2)\langle U^{\dagger 2} \rangle + \frac{(1 - p_A)(1 - p_{BPC})}{d^2} \\ &= p^3 + \frac{p(1 - p^2)}{d} \text{Tr}(U)\langle U^{\dagger} \rangle + p^2 \frac{(1 - p)}{d} \text{Tr}(U^2)\langle U^{\dagger 2} \rangle + \frac{(1 - p)(1 - p^2)}{d^2}, \\ R_{12} &= p_C + \frac{p_{APB}(1 - p_C)}{d} \text{Tr}(U^{\dagger})\langle U \rangle + \frac{(1 - p_A)p_B(1 - p_C)}{d^2} |\text{Tr}(U)|^2 + \frac{(1 - p_B)(1 - p_C)}{d^2} \\ &= p + \frac{p^2(1 - p)}{d} \text{Tr}(U^{\dagger})\langle U \rangle + \frac{(1 - p)^2 p}{d^2} |\text{Tr}(U)|^2 + \frac{(1 - p)^2}{d^2}, \end{aligned} \tag{5.3}$$

where we have kept distinct values of p_O on the first lines of the equations to show where the different terms originate; when the three channels are truly identical, we can set them each to be the same variable p . All three functions are linearly independent, even with maximally mixed probe states that will make these three into real functions that depend on $u(\theta)$, $u(2\theta)$, p , and d .

A measurement of the control system in the $P_{ij\pm}$ basis will thus yield information from which the unitary's phase and the depolarization noise's strength can simultaneously be estimated. One only needs to measure two of the off-diagonal components to find these two parameters, such as $(|0\rangle \pm |1\rangle)/\sqrt{2}$ and $(|0\rangle \pm |2\rangle)/\sqrt{2}$, but redundant information can be obtained by measuring the other components and one can also use these to simultaneously estimate the dimension parameter d if it is unknown.

These demonstrate how increasing the dimension of the control system gives access to multiparameter estimation techniques for ICO strategies with multiple copies of the same noisy unitaries. Similar results can straightforwardly be obtained for other noise channels, where one can also revisit the question of measuring more than one parameter of the unitary simultaneously with only maximally mixed probe states. Here, one can also learn about multiple properties of the unitary simultaneously by using a probe state other than the maximally mixed one, as there are three different complex functional dependencies of the R_{ij} on U when the control system dimension was simply increased from 2 to 3. Higher dimensional controls lead to more simultaneously estimable parameters.

What happens when a control system controls the order in which two copies of $\mathcal{E}_{U\text{-depol}}$ are applied to a probe? For $d = 2$, this has been studied in Ref. [71]. Rather than simply extend this result to arbitrary d , we also allow for three copies of the same channel, increasing the dimension of the control state to allow for multiple parameters to simultaneously be estimated. With control states $|0\rangle$, $|1\rangle$, and $|2\rangle$ dictating that the probe experiences the noisy unitary channels in orders $\mathcal{E}^{(A)} \circ \mathcal{E}^{(B)} \circ \mathcal{E}^{(C)}$, $\mathcal{E}^{(C)} \circ \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}$, and $\mathcal{E}^{(B)} \circ \mathcal{E}^{(C)} \circ \mathcal{E}^{(A)}$, respectively, the three off-diagonal matrix elements of the evolved control state ρ'_c require the three functions

Next, we consider the extension to D copies of the depolarization channel by allowing the control state $|0\rangle$ to dictate the order $\mathcal{E}^{(A_0)} \circ \mathcal{E}^{(A_1)} \circ \mathcal{E}^{(A_2)} \circ \dots \circ \mathcal{E}^{(A_{D-1})}$ and $|1\rangle$ to dictate $\mathcal{E}^{(A_1)} \circ \mathcal{E}^{(A_2)} \circ \dots \circ \mathcal{E}^{(A_{D-1})} \circ \mathcal{E}^{(A_0)}$. Keeping terms to lowest order in p means that we need only consider at most one of the channels to contribute the Kraus operator $K_{\mathbb{1}}$. These terms are identical when that single identity Kraus operator comes from any channel other than A_0 , so we can readily compute

$$R_{01} \approx \frac{1-p}{d^2} + \frac{p}{d} \text{Tr}(U)\langle U^{\dagger} \rangle. \tag{5.4}$$

The dependence on U is only diminished by $O(p)$, instead of by $O(p^D)$ for D passes through a depolarization channel, which constitutes another large advantage over naive DCO schemes, with the caveat that adaptive or ancilla-assisted schemes should be able to attain our scaling in the limit of unlimited copies of the noisy channel U .

VI. RESULTS INDEPENDENT FROM PROBE STATE: WHEN THE PROBE IS A QUBIT

In most of the examples above, the probe state was chosen to be the maximally mixed state to showcase the capabilities of ICO: ICO allows a maximally insensitive state to become sensitive. In a related but different context of ICO metrology with qubit probe states, it was found that the results were independent from the probe state [71,103], implying that maximally mixed states would achieve the same results as any other probe state. We show here how this trend can be generalized to ICO schemes with arbitrary numbers of channels,

including having multiple copies of U , multiple copies of the noise channels, and more than two different noise channels.

Arbitrary qubit states can be decomposed as

$$\rho_p = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma}), \quad (6.1)$$

so we desire to show that all of the parameter dependence can be imprinted onto the control system ρ'_c in a manner independent of \mathbf{r} . This is equivalent to showing that $R_{j_1 j_2}(\rho_p) = R_{j_1 j_2}(\mathbb{1}/2)$ or that $R_{j_1 j_2}(\sigma_i) = 0 \forall i \in (1, 2, 3)$. Actually, this property only holds true for some specific sequences of causal orders and some particular noisy channels. What we can instead prove is that

$$\text{Re}[R_{j_1 j_2}(\rho_p)] = \text{Re}[R_{j_1 j_2}(\mathbb{1}/2)] \Leftrightarrow \text{Re}[R_{j_1 j_2}(\sigma_i)] = 0 \quad (6.2)$$

$$\text{Re} \left[\text{Tr} \left(\sum_{i_1, \dots, i_3} K_{i_{\pi_{j_1}(0)}}^{(A_{\pi_{j_1}(0)})} \dots K_{i_{\pi_{j_1}(D-1)}}^{(A_{\pi_{j_1}(D-1)})} \boldsymbol{\sigma} (K_{i_{\pi_{j_2}(0)}}^{(A_{\pi_{j_2}(0)})} \dots K_{i_{\pi_{j_2}(D-1)}}^{(A_{\pi_{j_2}(D-1)})})^\dagger \right) \right] = \mathbf{0}, \quad \forall K_i^{(A_j)} = \alpha(i, j)U(i, j), \quad (6.4)$$

where each value of α and U can vary with i and j . Although this expression looks formidable, it can be proven using routine properties of Pauli matrices,

$$\sigma_\mu \sigma_\nu = \delta_{\mu\nu} \sigma_0 + i \sum_{\lambda=1}^3 \epsilon_{\mu\nu\lambda} \sigma_\lambda, \quad (6.5)$$

where we have used $\sigma_0 = \mathbb{1}$, the Kronecker delta $\delta_{\mu\nu}$, and the fully antisymmetric Levi-Civita tensor $\epsilon_{\mu\nu\lambda}$. The constants α are immaterial to the proof of Eq. (6.4) so we need not keep track of them. In fact, because each Kraus operator and its Hermitian conjugate appears in Eq. (6.4), the global phase of each Kraus operator is irrelevant, so we can always consider α to be real.

A single Kraus operator takes the form

$$K = \alpha \sigma_0 + i \mathbf{v} \cdot \boldsymbol{\sigma} \quad (6.6)$$

for some real vector $\mathbf{v} = (v_1, v_2, v_3)$ and real constant α . The product of two Kraus operators of the form of Eq. (6.6) is another Kraus operator of the same form:

$$\begin{aligned} (\alpha \sigma_0 + i \mathbf{v} \cdot \boldsymbol{\sigma})(\alpha' \sigma_0 + i \mathbf{v}' \cdot \boldsymbol{\sigma}) \\ = (\alpha \alpha' - \mathbf{v} \cdot \mathbf{v}') \sigma_0 + i(\alpha \mathbf{v}' + \alpha' \mathbf{v} - \mathbf{v} \times \mathbf{v}') \cdot \boldsymbol{\sigma} \\ \equiv \alpha'' \sigma_0 + i \mathbf{v}'' \cdot \boldsymbol{\sigma}. \end{aligned} \quad (6.7)$$

The important property is that α'' is still real, which is not true for a generic multiplication of two unitary operators. We therefore infer that

$$K_{i_{\pi_{j_1}(0)}}^{(A_{\pi_{j_1}(0)})} \dots K_{i_{\pi_{j_1}(D-1)}}^{(A_{\pi_{j_1}(D-1)})} = \beta \sigma_0 + i \mathbf{u} \cdot \boldsymbol{\sigma}, \quad (6.8)$$

for some real β and vector \mathbf{u} , and similarly for the Hermitian conjugates of these Kraus operators with another real β' and vector \mathbf{u}' .

We are now equipped to tackle the expression in Eq. (6.4). By the cyclic nature of the trace, we simply need to show that

$$\text{Re}\{\text{Tr}[(\beta'' \sigma_0 + i \mathbf{u}'' \cdot \boldsymbol{\sigma}) \boldsymbol{\sigma}]\} = \mathbf{0}. \quad (6.9)$$

for all channels with Kraus operators

$$K_i^{(A_j)} K_i^{(A_j)\dagger} = K_i^{(A_j)\dagger} K_i^{(A_j)} \propto \mathbb{1}. \quad (6.3)$$

That is, we show that measuring the real parts of the off-diagonal elements of ρ'_c , equivalent to measuring ρ'_c in the $(|i\rangle \pm |j\rangle)/\sqrt{2}$ basis when ρ_c is initialized with coefficients of equal phase, gives information about the channels that is independent of the probe state whenever each Kraus operator from the channels can be written as some constant multiplied some unitary. No relationships between any two Kraus operators $K_i^{(A_j)}$ and $K_i^{(A_{j'})}$ are necessary to enable our broad result.

As a consequence, we seek a proof that

Each of the components of $\boldsymbol{\sigma}$ is traceless, so $\text{Tr}(\sigma_0 \boldsymbol{\sigma}) = \mathbf{0}$. Similarly, using Eq. (6.5), we find that

$$\text{Tr}[(\mathbf{u}'' \cdot \boldsymbol{\sigma}) \boldsymbol{\sigma}] = 2 \mathbf{u}'', \quad (6.10)$$

where the factor of 2 comes from $\text{Tr}(\sigma_0) = 2$. The vector \mathbf{u}'' is always real, as explained above, which immediately proves Eq. (6.4).

How common are such Kraus operators that satisfy $K_i^\dagger K_i \propto \mathbb{1}$? This is manifestly satisfied by the dephasing channel, with Kraus operators proportional to $\mathbb{1}$ and σ_u . For the depolarizing channel, we need a Kraus-operator decomposition other than the one used above to show that it also satisfies this condition, remembering that different sets of Kraus operators can lead to identical dynamics if they are related by unitary transformations as $K'_i = \sum_j U_{ij} K_j$ with unitary matrices \mathbf{U} . A possible decomposition of the depolarizing channel is into the four Kraus operators $\frac{\sqrt{1+3p}}{2} \mathbb{1}$, $\frac{\sqrt{1-p}}{2} \sigma_1$, $\frac{\sqrt{1-p}}{2} \sigma_2$, and $\frac{\sqrt{1-p}}{2} \sigma_3$, which are manifestly proportional to unitary matrices. As for the amplitude damping channel, one can verify that matrix elements such as R_{01} do depend on the initial probe state, changing from its expression in Eq. (3.20) for maximally mixed probes to $R_{01} = p_A p_B$ for probe state $\rho_p = |1\rangle\langle 1|$, from which one can conclude that there is no Kraus-operator decomposition for amplitude damping channels in which the Kraus operators are proportional to unitary matrices. Any number of depolarization channels and dephasing channels, as well as their generalizations into Pauli channels [117, 118], supplied in any number of coherently controlled orders, will lead to control states the real part of whose elements will be independent from the probe state that traversed the channels.

VII. CONCLUDING REMARKS

ICO opens many doors to quantum-enhanced metrology. We showed how a variety of noise channels that would otherwise eradicate all hopes of measuring parameters could be circumvented by ICO to allow those parameters to be estimated, with dramatic scaling advantages over any causally ordered scheme. Our protocols only require measurement of

a control system that did not probe the unitary in question and allow one to simultaneously estimate multiple unitary and noise parameters. All of the protocols detailed here are readily accessible to experiments that have already investigated ICO using a quantum switch, especially those that studied communication through noisy channels with ICO. They are especially experimentally friendly due to the probe states being maximally mixed and the measurements being projections on superposition states standard to interferometry. We hope this incorporation of multiparameter estimation to ICO continues to be a fruitful breeding ground for many more quantum advantages.

The important distinction between this and previous works that studied metrology augmented by ICO are our allowance of the order between the noise and unitary channels to also be controlled; whereas previous works only had access to controlling the order of multiple copies of identical noisy operations, so they could never achieve the results of this paper in the limit of completely noisy channels. In the case of identical noisy operations and single-parameter estimation, landmark studies showed that ICO should always provide at least a small advantage [72], but such an advantage disappears in the asymptotic limit of infinite copies of the noisy operations, where adaptive and entangled-ancilla protocols are equally as effective as ICO [116]. These leave open intriguing questions. In terms of having multiple copies of identical channels, as in Sec. V, is there a hierarchy of estimation strategies for multiparameter estimation? Does ICO retain any advantage in the asymptotic limit for multiparameter estimation? And, for all of our findings in Secs. III and IV, where we allow for the order in which the noise channels and the unitary are applied to be controlled, how do the advantages of ICO evolve when multiple copies of the noise and unitary channels are allowed? With two identical unitaries and four noise channels, does the ICO advantage increase or decrease? In the asymptotic limit of a large number of identical unitaries

and noise channels, can rigorous inequalities or equalities be proven between different classes of estimation strategies? We know that controlling the order of a single unitary and a single completely depolarizing channel, as described here, will outperform even an infinite number of applications of identical unitary channels that are each always subject to complete depolarization noise; we thus expect ICO to have even greater advantages as the number of copies of the channels is increased.

ACKNOWLEDGMENTS

The authors are grateful for discussions with K. Bonsma-Fisher, F. Bouchard, D. England, K. Fenwick, B. Freitas, and B. Sussman. We also thank the International Network of Acausal Quantum Technology, funded by the Engineering and Physical Sciences Research Council (EPSRC), for support. A.Z.G. and K.H. acknowledge that the NRC headquarters is located on the traditional unceded territory of the Algonquin Anishinaabe and Mohawk people, as well as support from NRC’s Quantum Sensors Challenge Program. A.Z.G. acknowledges funding from the NSERC PDF program. L.L.S.S. acknowledges support from Ministerio de Ciencia e Innovación (Grant No. PID2021-127781NB-I00).

APPENDIX A: FISHER INFORMATION MATRIX FOR SIMULTANEOUS ESTIMATION OF DEPOLARIZATION NOISE AND UNITARY PHASE

In the main text, we calculated $P_{ij\pm}$ when the quantum switch controls three different orders of traversing the unitary channel and the two depolarization noise channels. The FI matrix can be calculated from these probabilities, which we list in turn, recalling that the matrix is always symmetric:

$$\begin{aligned}
 F_{\theta\theta} &= \frac{(\partial_u)^2}{6d^2} \left[\frac{p_A^2(p_B - 1)^2}{d^2(p_B + 1) - (p_B - 1)(p_A(u - 1) + 1)} + \frac{p_A^2(1 - p_B)}{d^2 - p_A u + p_A - 1} + \frac{(1 - p_A)^2 p_B^2}{d^2(p_A + 1) - (p_A - 1)(p_B(u - 1) + 1)} \right. \\
 &\quad \left. + \frac{(1 - p_A)p_B^2}{d^2 - p_B u + p_B - 1} + \frac{1 - p_A p_B}{d^2 - u} + \frac{(1 - p_A p_B)^2}{d^2(p_A p_B + 1) - p_A p_B u + u} \right], \\
 F_{\theta p_A} &= \frac{\partial_u}{6d^2} \left[-\frac{p_A(p_B - 1)^2(u - 1)}{(p_B - 1)(p_A(u - 1) + 1) - d^2(p_B + 1)} + \frac{p_A(p_B - 1)(u - 1)}{-d^2 + p_A(u - 1) + 1} \right. \\
 &\quad \left. + \frac{(p_A - 1)p_B(d^2 - p_B u + p_B - 1)}{(p_A - 1)(p_B(u - 1) + 1) - d^2(p_A + 1)} - \frac{p_B(d^2 - u)(p_A p_B - 1)}{d^2(p_A p_B + 1) - p_A p_B u + u} + 2p_B \right], \\
 F_{\theta p_B} &= \frac{\partial_u}{6d^2} \left[-\frac{(p_A - 1)^2 p_B(u - 1)}{(p_A - 1)(p_B(u - 1) + 1) - d^2(p_A + 1)} + \frac{(p_A - 1)p_B(u - 1)}{-d^2 + p_B(u - 1) + 1} \right. \\
 &\quad \left. + \frac{p_A(p_B - 1)(d^2 - p_A u + p_A - 1)}{(p_B - 1)(p_A(u - 1) + 1) - d^2(p_B + 1)} - \frac{p_A(d^2 - u)(p_A p_B - 1)}{d^2(p_A p_B + 1) - p_A p_B u + u} + 2p_A \right], \\
 F_{p_A p_A} &= \frac{1}{6d^2} \left[\frac{p_B^2(u - d^2)}{p_A p_B - 1} + \frac{p_B^2(d^2 - u)^2}{d^2(p_A p_B + 1) - p_A p_B u + u} - \frac{d^2 - p_B u + p_B - 1}{p_A - 1} - \frac{(p_B - 1)(u - 1)^2}{d^2 - p_A u + p_A - 1} \right. \\
 &\quad \left. + \frac{(p_B - 1)^2(u - 1)^2}{d^2(p_B + 1) - (p_B - 1)(p_A(u - 1) + 1)} + \frac{(d^2 - p_B u + p_B - 1)^2}{d^2(p_A + 1) - (p_A - 1)(p_B(u - 1) + 1)} \right],
 \end{aligned}$$

TABLE I. Matrix elements R_{ij} contributing to the updated control state ρ'_c after the probe traverses the channels in a particular order and is then ignored or discarded. The rows label the order of operations when the control is in state $|i\rangle$ and the columns when the control is in state $|j\rangle$. The other half of the combinations are tabulated in Table II.

	$ 0\rangle: \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)} \circ U$	$ 1\rangle: U \circ \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}$	$ 2\rangle: \mathcal{E}^{(B)} \circ U \circ \mathcal{E}^{(A)}$
$ 0\rangle: \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)} \circ U$	1	$p_A p_B + \frac{1-p_A p_B}{d} \tilde{u}$	$p_A + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_B \bar{p}_A}{d} \tilde{u}$
$ 1\rangle: U \circ \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}$	$p_A p_B + \frac{1-p_A p_B}{d} \tilde{u}^*$	1	$p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u}^*$
$ 2\rangle: \mathcal{E}^{(B)} \circ U \circ \mathcal{E}^{(A)}$	$p_A + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_B \bar{p}_A}{d} \tilde{u}^*$	$p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u}$	1
$ 3\rangle: U \circ \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B + p_B \bar{p}_A}{d} \tilde{u}^*$	$p_A + p_B - p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2}$	$p_B + \frac{p_A \bar{p}_B}{d} \tilde{u}^* + \frac{\bar{p}_A \bar{p}_B}{d^2} u$
$ 4\rangle: \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)} \circ U$	$p_A + p_B - p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2}$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B + p_B \bar{p}_A}{d} \tilde{u}$	$p_A + \frac{1-p_A}{d} \tilde{u}$
$ 5\rangle: \mathcal{E}^{(A)} \circ U \circ \mathcal{E}^{(B)}$	$p_B + \frac{1-p_B}{d} \tilde{u}^*$	$p_A + \frac{p_B \bar{p}_A}{d} \tilde{u} + \frac{\bar{p}_A \bar{p}_B}{d^2} u$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u}^* + \frac{p_B \bar{p}_A}{d} \tilde{u}$

$$\begin{aligned}
 F_{p_A p_B} &= \frac{1}{6d^2} \left\{ 2u - 2 + \frac{p_A p_B (u - d^2)}{p_A p_B - 1} + \frac{p_A p_B (d^2 - u)^2}{d^2 (p_A p_B + 1) - p_A p_B u + u} \right. \\
 &\quad \left. + (2u - 2) \left[\frac{d^2}{d^2 (p_B + 1) - (p_B - 1)(p_A (u - 1) + 1)} + \frac{d^2}{d^2 (p_A + 1) - (p_A - 1)(p_B (u - 1) + 1)} - 1 \right] \right\}, \\
 F_{p_B p_B} &= \frac{1}{6} \left[\frac{p_A^2 (u - d^2)}{d^2 (p_A p_B - 1)} + \frac{p_A^2 (d^2 - u)^2}{d^2 (d^2 (p_A p_B + 1) - p_A p_B u + u)} - \frac{d^2 - p_A u + p_A - 1}{d^2 (p_B - 1)} \right. \\
 &\quad \left. - \frac{(p_A - 1)(u - 1)^2}{d^2 (d^2 - p_B u + p_B - 1)} + \frac{(p_A - 1)^2 (u - 1)^2}{d^4 (p_A + 1) - d^2 (p_A - 1)(p_B (u - 1) + 1)} + \frac{(d^2 - p_A u + p_A - 1)^2}{d^4 (p_B + 1) - d^2 (p_B - 1)(p_A (u - 1) + 1)} \right]. \tag{A1}
 \end{aligned}$$

APPENDIX B: ALL OFF-DIAGONAL ELEMENTS FOR TWO DEPOLARIZATION CHANNELS AND ONE UNITARY CHANNEL

There are six orders in which a probe system can traverse three channels. We compute the relevant coefficients R_{ij} for all interference terms between all 36 pairs of channels that arise in Eqs. (2.12) and (2.14). We define $\tilde{u} = \langle U \rangle \text{Tr}(U^\dagger)$ for expectation values with respect to the initial probe state and $u = |\text{Tr}(U)|^2$ as above. To conserve space, we define $\bar{p}_O = 1 - p_O$ and split the 36 matrix elements into two tables, Tables I and II.

APPENDIX C: FISHER INFORMATION MATRIX FOR SIMULTANEOUS ESTIMATION OF DEPHASING NOISE AND UNITARY PHASE

In the main text, we calculated $P_{ij\pm}$ when the quantum switch controls three different orders of traversing the unitary channel and the two dephasing noise channels. The FI matrix can be calculated from these probabilities, which we list in turn, recalling that the matrix is always symmetric:

$$\begin{aligned}
 F_{\theta\theta} &= \frac{1}{3} \left[\frac{p_A (4p_B - 2) - 2p_B + 2}{(p_A (2p_B - 1) - p_B) \cos \theta - 2p_A p_B + p_A + p_B - 2} - \frac{2p_A}{p_A (-\cos \theta) + p_A + \cos \theta + 1} \right. \\
 &\quad \left. - \frac{2p_B}{p_B (-\cos \theta) + p_B + \cos \theta + 1} + 3 \right],
 \end{aligned}$$

TABLE II. Matrix elements R_{ij} contributing to the updated control state ρ'_c after the probe traverses the channels in a particular order and is then ignored or discarded. The rows label the order of operations when the control is in state $|i\rangle$ and the columns when the control is in state $|j\rangle$. The other half of the combinations are tabulated in Table I.

	$ 3\rangle: U \circ \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}$	$ 4\rangle: \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)} \circ U$	$ 5\rangle: \mathcal{E}^{(A)} \circ U \circ \mathcal{E}^{(B)}$
$ 0\rangle: \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)} \circ U$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B + p_B \bar{p}_A}{d} \tilde{u}$	$p_A + p_B - p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2}$	$p_B + \frac{1-p_B}{d} \tilde{u}$
$ 1\rangle: U \circ \mathcal{E}^{(A)} \circ \mathcal{E}^{(B)}$	$p_A + p_B - p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2}$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B + p_B \bar{p}_A}{d} \tilde{u}^*$	$p_A + \frac{p_B \bar{p}_A}{d} \tilde{u}^* + \frac{\bar{p}_A \bar{p}_B}{d^2} u$
$ 2\rangle: \mathcal{E}^{(B)} \circ U \circ \mathcal{E}^{(A)}$	$p_B + \frac{p_A \bar{p}_B}{d} \tilde{u} + \frac{\bar{p}_A \bar{p}_B}{d^2} u$	$p_A + \frac{1-p_A}{d} \tilde{u}^*$	$p_A p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u} + \frac{p_B \bar{p}_A}{d} \tilde{u}^*$
$ 3\rangle: U \circ \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)}$	1	$p_A p_B + \frac{1-p_A p_B}{d} \tilde{u}^*$	$p_A + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_B \bar{p}_A}{d} \tilde{u}^*$
$ 4\rangle: \mathcal{E}^{(B)} \circ \mathcal{E}^{(A)} \circ U$	$p_A p_B + \frac{1-p_A p_B}{d} \tilde{u}$	1	$p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u}$
$ 5\rangle: \mathcal{E}^{(A)} \circ U \circ \mathcal{E}^{(B)}$	$p_A + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_B \bar{p}_A}{d} \tilde{u}$	$p_B + \frac{\bar{p}_A \bar{p}_B}{d^2} + \frac{p_A \bar{p}_B}{d} \tilde{u}^*$	1

$$\begin{aligned}
F_{\theta p_A} &= \frac{1}{3} \sin \theta \left[\frac{2p_B - 1}{(p_A(2p_B - 1) - p_B) \cos \theta - 2p_A p_B + p_A + p_B - 2} - \frac{1}{p_A(-\cos \theta) + p_A + \cos \theta + 1} \right], \\
F_{\theta p_B} &= \frac{1}{3} \sin \theta \left[\frac{2p_A - 1}{(p_A(2p_B - 1) - p_B) \cos \theta - 2p_A p_B + p_A + p_B - 2} - \frac{1}{p_B(-\cos \theta) + p_B + \cos \theta + 1} \right], \\
F_{p_A p_B} &= \frac{1}{3} \left[\frac{(1 - 2p_B)^2 \sin^2 \frac{\theta}{2}}{-2p_A p_B + p_A + p_B} + \frac{(1 - 2p_B)^2 (\cos \theta - 1)^2}{2(-2p_A p_B + p_A + p_B) \cos \theta + p_A(4p_B - 2) - 2p_B + 4} \right. \\
&\quad \left. + \frac{(\cos \theta - 1)^2}{2(p_A(-\cos \theta) + p_A + \cos \theta + 1)} + \frac{\cos \theta - 1}{2p_A - 2} \right], \\
F_{p_A} &= \frac{(2p_A - 1)(2p_B - 1)(\cos \theta - 1)}{3(-2p_A p_B + p_A + p_B)((p_A(2p_B - 1) - p_B) \cos \theta - 2p_A p_B + p_A + p_B - 2)}, \\
F_{p_B} &= \frac{1}{3} \left[\frac{(1 - 2p_A)^2 \sin^2 \frac{\theta}{2}}{-2p_A p_B + p_A + p_B} + \frac{(1 - 2p_A)^2 (\cos \theta - 1)^2}{2(-2p_A p_B + p_A + p_B) \cos \theta + p_A(4p_B - 2) - 2p_B + 4} \right. \\
&\quad \left. + \frac{(\cos \theta - 1)^2}{2(-p_B \cos \theta + p_B + \cos \theta + 1)} + \frac{\cos \theta - 1}{2p_B - 2} \right]. \tag{C1}
\end{aligned}$$

-
- [1] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, *Nat. Phys.* **7**, 406 (2011).
- [2] C. M. Caves, Quantum-mechanical noise in an interferometer, *Phys. Rev. D* **23**, 1693 (1981).
- [3] J. P. Dowling, Correlated input-port, matter-wave interferometer: Quantum-noise limits to the atom-laser gyroscope, *Phys. Rev. A* **57**, 4736 (1998).
- [4] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, *Science* **306**, 1330 (2004).
- [5] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, How to perform the most accurate possible phase measurements, *Phys. Rev. A* **80**, 052114 (2009).
- [6] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, Biological measurement beyond the quantum limit, *Nat. Photon.* **7**, 229 (2013).
- [7] M. Tsang, R. Nair, and X.-M. Lu, Quantum Theory of Superresolution for Two Incoherent Optical Point Sources, *Phys. Rev. X* **6**, 031033 (2016).
- [8] L.-Z. Liu, Y.-Z. Zhang, Z.-D. Li, R. Zhang, X.-F. Yin, Y.-Y. Fei, L. Li, N.-L. Liu, F. Xu, Y.-A. Chen, and J.-W. Pan, Distributed quantum phase estimation with entangled photons, *Nat. Photon.* **15**, 137 (2021).
- [9] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Super-resolving phase measurements with a multiphoton entangled state, *Nature (London)* **429**, 161 (2004).
- [10] The LIGO Scientific Collaboration, A gravitational wave observatory operating beyond the quantum shot-noise limit, *Nat. Phys.* **7**, 962 (2011).
- [11] R. Whittaker, C. Erven, A. Neville, M. Berry, J. L. O'Brien, H. Cable, and J. C. F. Matthews, Absorption spectroscopy at the ultimate quantum limit from single-photon states, *New J. Phys.* **19**, 023013 (2017).
- [12] C. You, M. Hong, P. Bierhorst, A. E. Lita, S. Glancy, S. Kolthammer, E. Knill, S. W. Nam, R. P. Mirin, O. S. Magaña-Loaiza, and T. Gerrits, Scalable multiphoton quantum metrology with neither pre- nor post-selected measurements, *Appl. Phys. Rev.* **8**, 041406 (2021).
- [13] J. Qin, Y.-H. Deng, H.-S. Zhong, L.-C. Peng, H. Su, Y.-H. Luo, J.-M. Xu, D. Wu, S.-Q. Gong, H.-L. Liu, H. Wang, M.-C. Chen, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Unconditional and Robust Quantum Metrological Advantage Beyond NOON States, *Phys. Rev. Lett.* **130**, 070801 (2023).
- [14] M. G. Raymer and C. Monroe, The US national quantum initiative, *Quantum Sci. Technol.* **4**, 020504 (2019).
- [15] B. Sussman, P. Corkum, A. Blais, D. Cory, and A. Damascelli, Quantum Canada, *Quantum Sci. Technol.* **4**, 020503 (2019).
- [16] Y. Yamamoto, M. Sasaki, and H. Takesue, Quantum information science and technology in Japan, *Quantum Sci. Technol.* **4**, 020502 (2019).
- [17] OIDA, Oida quantum photonics roadmap: Every photon counts, *OIDA Optica Industry Report*, 3 (2020).
- [18] P. Knight and I. Walmsley, UK national quantum technology programme, *Quantum Sci. Technol.* **4**, 040502 (2019).
- [19] K. Matsumoto, A new approach to the Cramér-Rao-type bound of the pure-state model, *J. Phys. A: Math. Theor.* **35**, 3111 (2002).
- [20] M. G. A. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inform.* **07**, 125 (2009).
- [21] G. Tóth and I. Apellaniz, Quantum metrology from a quantum information science perspective, *J. Phys. A: Math. Theor.* **47**, 424006 (2014).
- [22] M. Szczykulska, T. Baumgratz, and A. Datta, Multi-parameter quantum metrology, *Adv. Phys. X* **1**, 621 (2016).
- [23] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Quantum-enhanced measurements without entanglement, *Rev. Mod. Phys.* **90**, 035006 (2018).

- [24] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, *J. Phys. A: Math. Theor.* **53**, 023001 (2020).
- [25] J. S. Sidhu and P. Kok, Geometric perspective on quantum parameter estimation, *AVS Quantum Sci.* **2**, 014701 (2020).
- [26] F. Albarelli, M. Barbieri, M. Genoni, and I. Gianani, A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging, *Phys. Lett. A* **384**, 126311 (2020).
- [27] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, *AVS Quantum Sci.* **2**, 024703 (2020).
- [28] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, Multiparameter estimation beyond quantum Fisher information, *J. Phys. A: Math. Theor.* **53**, 363001 (2020).
- [29] J. Liu, M. Zhang, H. Chen, L. Wang, and H. Yuan, Optimal scheme for quantum metrology, *Adv. Quantum Technol.* **5**, 2100080 (2022).
- [30] A. Z. Goldberg, J. L. Romero, A. S. Sanz, and L. L. Sánchez-Soto, Taming singularities of the quantum Fisher information, *Int. J. Quantum Inform.* **19**, 2140004 (2021).
- [31] H. Zhu, Information complementarity: A new paradigm for decoding quantum incompatibility, *Sci. Rep.* **5**, 14317 (2015).
- [32] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Compatibility in multiparameter quantum metrology, *Phys. Rev. A* **94**, 052108 (2016).
- [33] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, *J. Phys. A: Math. Theor.* **49**, 123001 (2016).
- [34] F. Albarelli, J. F. Friel, and A. Datta, Evaluating the Holevo Cramér-Rao Bound for Multiparameter Quantum Metrology, *Phys. Rev. Lett.* **123**, 200503 (2019).
- [35] F. Belliardo and V. Giovannetti, Incompatibility in quantum parameter estimation, *New J. Phys.* **23**, 063055 (2021).
- [36] J. Suzuki, Nuisance parameter problem in quantum estimation theory: Tradeoff relation and qubit examples, *J. Phys. A: Math. Theor.* **53**, 264001 (2020).
- [37] J. Suzuki, Y. Yang, and M. Hayashi, Quantum state estimation with nuisance parameters, *J. Phys. A: Math. Theor.* **53**, 453001 (2020).
- [38] A. Z. Goldberg, L. L. Sánchez-Soto, and H. Ferretti, Intrinsic Sensitivity Limits for Multiparameter Quantum Metrology, *Phys. Rev. Lett.* **127**, 110501 (2021).
- [39] M. Koschorreck, M. Napolitano, B. Dubost, and M. W. Mitchell, High resolution magnetic vector-field imaging with cold atomic ensembles, *Appl. Phys. Lett.* **98**, 074101 (2011).
- [40] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Quantum Enhanced Multiple Phase Estimation, *Phys. Rev. Lett.* **111**, 070403 (2013).
- [41] T. Baumgratz and A. Datta, Quantum Enhanced Estimation of a Multidimensional Field, *Phys. Rev. Lett.* **116**, 030801 (2016).
- [42] J. Řehaček, Z. Hradil, B. Stoklasa, M. Paúr, J. Grover, A. Krzic, and L. L. Sánchez-Soto, Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution, *Phys. Rev. A* **96**, 062107 (2017).
- [43] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multiparameter Estimation in Networked Quantum Sensors, *Phys. Rev. Lett.* **120**, 080501 (2018).
- [44] J. Rubio, P. A. Knott, T. J. Proctor, and J. A. Dunningham, Quantum sensing networks for the estimation of linear functions, *J. Phys. A: Math. Theor.* **53**, 344001 (2020).
- [45] A. Z. Goldberg, A. B. Klimov, G. Leuchs, and L. L. Sánchez-Soto, Rotation sensing at the ultimate limit, *J. Phys. Photonics* **3**, 022008 (2021).
- [46] L. J. Fiderer, T. Tufarelli, S. Piano, and G. Adesso, General expressions for the quantum Fisher information matrix with applications to discrete quantum imaging, *PRX Quantum* **2**, 020308 (2021).
- [47] L. Hardy, Towards quantum gravity: A framework for probabilistic theories with non-fixed causal structure, *J. Phys. A: Math. Theor.* **40**, 3081 (2007).
- [48] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, *Phys. Rev. A* **88**, 022318 (2013).
- [49] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti, Quantum computation with programmable connections between gates, *Phys. Lett. A* **376**, 2940 (2012).
- [50] T. Morimae, Acausal measurement-based quantum computing, *Phys. Rev. A* **90**, 010101(R) (2014).
- [51] M. Araújo, F. Costa, and Č. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, *Phys. Rev. Lett.* **113**, 250402 (2014).
- [52] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, Computational advantage from the quantum superposition of multiple temporal orders of photonic gates, *PRX Quantum* **2**, 010320 (2021).
- [53] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, Quantum circuits with classical versus quantum control of causal order, *PRX Quantum* **2**, 030335 (2021).
- [54] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, *Phys. Rev. A* **86**, 040301(R) (2012).
- [55] A. Feix, M. Araújo, and Č. Brukner, Quantum superposition of the order of parties as a communication resource, *Phys. Rev. A* **92**, 052326 (2015).
- [56] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication, *Phys. Rev. Lett.* **117**, 100502 (2016).
- [57] F. Del Santo and B. Dakić, Two-Way Communication with a Single Quantum Particle, *Phys. Rev. Lett.* **120**, 060503 (2018).
- [58] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, *Phys. Rev. Lett.* **120**, 120502 (2018).
- [59] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, Communication enhancement through quantum coherent control of N channels in an indefinite causal-order scenario, *Entropy* **21**, 1012 (2019).
- [60] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, Indefinite causal order enables perfect quantum communication with zero capacity channels, *New J. Phys.* **23**, 033039 (2021).
- [61] D. Felce and V. Vedral, Quantum Refrigeration with Indefinite Causal Order, *Phys. Rev. Lett.* **125**, 070603 (2020).
- [62] H. Cao, N.-N. Wang, Z. Jia, C. Zhang, Y. Guo, B.-H. Liu, Y.-F. Huang, C.-F. Li, and G.-C. Guo, Quantum simulation of

- indefinite causal order induced quantum refrigeration, *Phys. Rev. Res.* **4**, L032029 (2022).
- [63] A. Z. Goldberg and K. Heshami, Breaking the limits of purification: Postselection enhances heat-bath algorithmic cooling, *J. Phys. Commun.* **7**, 015003 (2023).
- [64] H. Nie, T. Feng, S. Longden, and V. Vedral, Quantum cooling activated by coherent-controlled thermalisation, [arXiv:2201.06954](https://arxiv.org/abs/2201.06954).
- [65] T. Guha, M. Alimuddin, and P. Parashar, Thermodynamic advancement in the causally inseparable occurrence of thermal maps, *Phys. Rev. A* **102**, 032215 (2020).
- [66] T. Guha, S. Roy, K. Simonov, and Z. Zimborás, Activation of thermal states by quantum switch-driven thermalization and its limits, [arXiv:2208.04034](https://arxiv.org/abs/2208.04034).
- [67] K. Simonov, G. Francica, G. Guarnieri, and M. Paternostro, Work extraction from coherently activated maps via quantum switch, *Phys. Rev. A* **105**, 032217 (2022).
- [68] C. Mukhopadhyay, M. K. Gupta, and A. K. Pati, Superposition of causal order as a metrological resource for quantum thermometry, [arXiv:1812.07508](https://arxiv.org/abs/1812.07508).
- [69] M. Frey, Indefinite causal order aids quantum depolarizing channel identification, *Quantum Inf. Process.* **18**, 96 (2019).
- [70] X. Zhao, Y. Yang, and G. Chiribella, Quantum Metrology with Indefinite Causal Order, *Phys. Rev. Lett.* **124**, 190503 (2020).
- [71] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, *Phys. Rev. A* **103**, 032615 (2021).
- [72] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal Strategies of Quantum Metrology with a Strict Hierarchy, *Phys. Rev. Lett.* **130**, 070803 (2023).
- [73] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, *Nat. Commun.* **6**, 7913 (2015).
- [74] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, *Sci. Adv.* **3**, e1602589 (2017).
- [75] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, Experimental entanglement of temporal order, *Quantum* **6**, 621 (2022).
- [76] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, *Phys. Rev. Lett.* **121**, 090503 (2018).
- [77] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity, *Phys. Rev. Lett.* **122**, 120504 (2019).
- [78] F. Massa, A. Moqanaki, Ä. Baumeler, F. Del Santo, J. A. Kettlewell, B. Dakić, and P. Walther, Experimental two-way communication with one photon, *Adv. Quantum Technol.* **2**, 1900050 (2019).
- [79] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Increasing communication capacity via superposition of order, *Phys. Rev. Res.* **2**, 033292 (2020).
- [80] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, Experimental Transmission of Quantum Information Using a Superposition of Causal Orders, *Phys. Rev. Lett.* **124**, 030502 (2020).
- [81] D. Felce, V. Vedral, and F. Tennie, Refrigeration with indefinite causal orders on a cloud quantum computer, [arXiv:2107.12413](https://arxiv.org/abs/2107.12413).
- [82] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, Č. Brukner, A. A. Abbott, C. Branciard, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, *Phys. Rev. Res.* **3**, 013093 (2021).
- [83] P. Yin, X. Zhao, Y. Yang, Y. Guo, W.-H. Zhang, G.-C. Li, Y.-J. Han, B.-H. Liu, J.-S. Xu, G. Chiribella *et al.*, Experimental super-Heisenberg quantum metrology with indefinite gate order, *Nat. Phys.* **19**, 1122 (2023).
- [84] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012).
- [85] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, *New J. Phys.* **17**, 102001 (2015).
- [86] Ä. Baumeler and S. Wolf, The space of logically consistent classical processes without causal order, *New J. Phys.* **18**, 013036 (2016).
- [87] O. Oreshkov and C. Giarmatzi, Causal and causally separable processes, *New J. Phys.* **18**, 093020 (2016).
- [88] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, Bell's theorem for temporal order, *Nat. Commun.* **10**, 3772 (2019).
- [89] O. Oreshkov, Time-delocalized quantum subsystems and operations: On the existence of processes with indefinite causal structure in quantum mechanics, *Quantum* **3**, 206 (2019).
- [90] A. Dimić, M. Milivojević, D. Gočanin, N. S. Möller, and Č. Brukner, Simulating indefinite causal order with Rindler observers, *Front. Phys.* **8**, 470 (2020).
- [91] S. Milz, D. Jurkschat, F. A. Pollock, and K. Modi, Delayed-choice causal order and nonclassical correlations, *Phys. Rev. Res.* **3**, 023028 (2021).
- [92] J. Barrett, R. Lorenz, and O. Oreshkov, Cyclic quantum causal models, *Nat. Commun.* **12**, 885 (2021).
- [93] T. Purves and A. J. Short, Quantum Theory Cannot Violate a Causal Inequality, *Phys. Rev. Lett.* **127**, 110402 (2021).
- [94] V. Vilasini and R. Renner, Embedding cyclic causal structures in acyclic spacetimes: No-go results for process matrices, [arXiv:2203.11245](https://arxiv.org/abs/2203.11245).
- [95] F. Chapeau-Blondeau, Quantum parameter estimation on coherently superposed noisy channels, *Phys. Rev. A* **104**, 032214 (2021).
- [96] A. Goldberg, L. L. Sánchez-Soto, and K. Heshami, Measuring impossible parameters with indefinite causal order (unpublished).
- [97] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Quantum phase estimation with lossy interferometers, *Phys. Rev. A* **80**, 013825 (2009).
- [98] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, *Nat. Commun.* **3**, 1063 (2012).
- [99] P. J. D. Crowley, A. Datta, M. Barbieri, and I. A. Walmsley, Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry, *Phys. Rev. A* **89**, 023845 (2014).

- [100] I. Gianani, F. Albarelli, A. Verna, V. Cimini, R. Demkowicz-Dobrzanski, and M. Barbieri, Kramers–Kronig relations and precision limits in quantum phase estimation, *Optica* **8**, 1642 (2021).
- [101] S.-Y. Bai and J.-H. An, Floquet Engineering to Overcome no-go Theorem of Noisy Quantum Metrology, *Phys. Rev. Lett.* **131**, 050801 (2023).
- [102] G. Chiribella and X. Zhao, Heisenberg-limited metrology with coherent control on the probes’ configuration, [arXiv:2206.03052](https://arxiv.org/abs/2206.03052).
- [103] F. Chapeau-Blondeau, Indefinite causal order for quantum metrology with quantum thermal noise, *Phys. Lett. A* **447**, 128300 (2022).
- [104] F. Delgado, Symmetries of quantum Fisher information as parameter estimator for Pauli channels under indefinite causal order, *Symmetry* **14**, 1813 (2022).
- [105] If there are fewer channels, some of the Kraus operators can be set to an appropriate multiple of the identity and, if there are more channels, some of them can be concatenated into one channel [ABC and CAB to $(AB)C$ and $C(AB)$] or applied with a definite causal order (ABC and CBA always have operation B at the same time).
- [106] A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, On quantumness in multi-parameter quantum estimation, *J. Stat. Mech.* (2019) 094010.
- [107] M. Tsang, F. Albarelli, and A. Datta, Quantum Semiparametric Estimation, *Phys. Rev. X* **10**, 031023 (2020).
- [108] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [109] J. P. Dowling, Quantum optical metrology—the lowdown on high-NOON states, *Contemp. Phys.* **49**, 125 (2008).
- [110] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Bell’s theorem without inequalities, *Am. J. Phys.* **58**, 1131 (1990).
- [111] We take this opportunity to correct a typo in the factor of 2 that does not appear in earlier sources [119,120] but propagates in review articles [21,25].
- [112] N. Yunger Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, *Phys. Rev. A* **97**, 042105 (2018).
- [113] N. Gao, D. Li, A. Mishra, J. Yan, K. Simonov, and G. Chiribella, Measuring Incompatibility and Clustering Quantum Observables with a Quantum Switch, *Phys. Rev. Lett.* **130**, 170201 (2023).
- [114] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [115] See Ref. [72] for a comparison of ICO strategies for the amplitude damping channel when given access to two (or more) copies of U , each followed by an amplitude damping channel, where one can control the orders of the two noisy channels but cannot control the order of the noise and the unitary within each channel.
- [116] S. Kurdzialek, W. Gorecki, F. Albarelli, and R. Demkowicz-Dobrzanski, Using adaptiveness and causal superpositions against noise in quantum metrology, *Phys. Rev. Lett.* **131**, 090801 (2023).
- [117] T. S. Cubitt, M. B. Ruskai, and G. Smith, The structure of degradable quantum channels, *J. Math. Phys.* **49**, 102104 (2008).
- [118] D. Leung and J. Watrous, On the complementary quantum capacity of the depolarizing channel, *Quantum* **1**, 28 (2017).
- [119] G. Tóth, Multipartite entanglement and high-precision metrology, *Phys. Rev. A* **85**, 022322 (2012).
- [120] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, Fisher information and multiparticle entanglement, *Phys. Rev. A* **85**, 022321 (2012).