



**HAL**  
open science

# Reassessing the advantage of indefinite causal orders for quantum metrology

Raphaël Mothe, Cyril Branciard, Alastair A. Abbott

► **To cite this version:**

Raphaël Mothe, Cyril Branciard, Alastair A. Abbott. Reassessing the advantage of indefinite causal orders for quantum metrology. 2023. hal-04360267

**HAL Id: hal-04360267**

**<https://inria.hal.science/hal-04360267>**

Preprint submitted on 21 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Reassessing the advantage of indefinite causal orders for quantum metrology

Raphaël Mothe,<sup>1,2,\*</sup> Cyril Branciard,<sup>1</sup> and Alastair A. Abbott<sup>2</sup>

<sup>1</sup>*Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

<sup>2</sup>*Univ. Grenoble Alpes, Inria, 38000 Grenoble, France*

(Dated: December 20, 2023)

The quantum switch, the canonical example of a process with indefinite causal order, has been claimed to provide various advantages over processes with definite causal orders for some particular tasks in the field of quantum metrology. In this work, we argue that some of these advantages in fact do not hold if a fairer comparison is made. To this end, we consider a framework that allows for a proper comparison between the performance, quantified by the quantum Fisher information, of different classes of indefinite causal order processes and that of causal strategies on a given metrological task. More generally, by considering the recently proposed classes of circuits with classical or quantum control of the causal order, we come up with different examples where processes with indefinite causal order offer (or not) an advantage over processes with definite causal order, qualifying the interest of indefinite causal order regarding quantum metrology. As it turns out, for a range of examples, the class of quantum circuits with quantum control of causal order, which are known to be physically realizable, is shown to provide a strict advantage over causally ordered quantum circuits as well as over the class of quantum circuits with causal superposition. Thus, considering this class provides new evidence that indefinite causal order strategies can strictly outperform definite causal order strategies in quantum metrology. Moreover, it shows that the so-called dynamical control of causal order, a feature of quantum circuits with quantum control of the causal order but not of quantum circuits with mere causal superposition, can be a useful resource in quantum metrology.

## I. INTRODUCTION

In the usual formalism of quantum mechanics, the causal order between different operations, i.e., the order in which the operations are applied to a target system, is not subject to any kind of quantum indefiniteness. Yet, an extension of quantum mechanical laws encapsulated in the process matrix formalism was proposed in Ref. [1], which allows for processes with so-called indefinite causal order. For instance, by controlling the causal order between two operations with a quantum degree of freedom placed in a superposition, it is possible to generate superpositions of—and thus indefinite—causal orders. This phenomenon is at the heart of the canonical example of a process with indefinite causal order, the quantum switch [2]. In addition to the fundamental questions raised regarding quantum causality [1, 3, 4], these indefinite causal order processes have attracted attention through the various advantages they can provide in different subfields of quantum information over processes with well-defined causal orders [5–10].

Quantum metrology is one such field, with several studies examining the potential for the quantum switch to provide advantages over processes with well-defined causal orders [11–23]. A basic scenario in quantum metrology is that of channel parameter estimation, where the goal is to maximize the precision of an estimate of an unknown parameter of a quantum channel, given  $N$  uses (or copies) of the channel and with access to different kinds of quantum resources such as quantum superposition or entanglement. Thanks to the quantum Cramér-Rao bound [24, 25], one can quantify the possible precision of an estimate using the quantum Fisher information (QFI).

Several studies over the last few years have attempted to show that the quantum switch can provide advantages in this simplest metrological scenario of estimating the parameter of rather natural families of quantum channels, including depolarizing channels [12] and thermalizing channels [11], amongst others [14, 16–18, 23]. These works all show that using the quantum switch to sequentially apply these channels in a superposition of different orders provides an advantage over the simple sequential, causally definite, composition of the channels, and conclude, perhaps questionably, that indefinite causal order is a resource for quantum metrology. Indeed, one may wonder whether the quantum switch indeed outperforms any standard adaptive, fixed-order strategy in these tasks. This motivates us to analyze here this question in more detail and, more generally, to investigate the potential advantages of different types of causally indefinite strategies in quantum metrology [19]. We note that another result [13], which was recently demonstrated experimentally [26], showed an interesting and rigorous advantage using the quantum switch in a rather different kind of (multi-parameter) metrological task; here, however, we focus on the simple framework described above and studied in Refs. [11, 12, 14, 16–18, 23].

Rather than studying the performance of particular processes or strategies, such as the quantum switch or sequential composition, we look here to understand the capabilities of different families or classes of processes. We thereby build on the framework on Refs. [15, 19] to optimize the QFI, and thus bound the precision of estimation, over several such pertinent strategies. Indeed, Ref. [19] already considered two classes that include indefinite causal order: quantum circuits with causal superposition (QC-Sup), which includes the quantum switch as well as some other circuits with indefinite causal order; and the whole general class of process matrices, which includes all valid indefinite causal order processes (Gen). By comparing these classes to standard sequential

---

\* Corresponding author:  
raphael.mothe@neel.cnrs.fr

quantum circuits (quantum circuits with fixed causal order; QC-FO), and parallel circuits with entanglement (QC-Par), they established hierarchies of performance in estimating certain families of channels. The particular interest devoted to the quantum switch in metrological settings is due to its position as one of the few causally indefinite processes with a clear physical interpretation and potential realization [27–33]. It is known, however, that physically meaningful processes beyond the quantum switch, and more generally beyond QC-Sup, exist [34]; indeed, there are also causally indefinite processes that go beyond QC-FO by exploiting dynamical causal orders.

Here, we seek to study the physical limits of both causally definite and indefinite quantum metrology. We thereby consider two further classes of processes introduced in Ref. [34]: quantum circuits with classical control of the causal order (QC-CC), and the quantum counterpart of QC-CCs, namely quantum circuits with quantum control of the causal order (QC-QC). We show how the framework of Ref. [19] can be adapted to these classes to bound—in some cases tightly—the QFI, and thus metrological performance, in the metrological tasks considered. Revisiting the results of Refs. [11, 12], we find the claims therein to be unfounded, with no advantage attainable with processes exhibiting indefinite causal orders. Focusing on the  $N = 2, 3$  cases, we detail the different possible hierarchies and show that the potential advantage of indefinite causal order strategies crucially depends on the family of quantum channels whose parameter is being estimated. Several situations where indefinite causal order strategies outperform causally definite strategies could be identified, showing that indefinite causal order can indeed be seen as a resource in quantum metrology. A strict advantage of QC-QCs over QC-Sups could also be identified for various situations, showing that the dynamical control of the causal order can also be seen as a quantum metrological resource and probing, in some specific tasks, the underlying physical limits of quantum metrology.

This paper is organized as follows. In Sec. II we give a brief overview of quantum metrology and causally indefinite matrices using the process matrix framework. In Sec. III we study in more detail the purported advantages of Refs. [11, 12] and show that the quantum switch can be outperformed by causally definite strategies. In Sec. IV we show how the metrological task can be solved using semidefinite programming over different classes of strategies, which we detail in Sec. V. In Sec. VI we present a comparison of the performance of different classes of strategies when estimating different channel families' parameters, before presenting our conclusions in Sec. VII.

## II. PRELIMINARIES

### A. Quantum metrology

In its simplest version, with one use of the channel, the metrological problem of channel parameter estimation can be described as follows (see Fig. 1). Consider a quantum channel  $\mathcal{C}_\theta$  parametrized (in a smooth enough manner) by an unknown continuous parameter  $\theta$  that we wish to estimate. We can probe

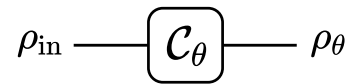


Figure 1. The basic typical scenario of channel parameter estimation. A probe state  $\rho_{\text{in}}$  is sent through a quantum channel  $\mathcal{C}_\theta$  specified by an unknown parameter  $\theta$ . The parameter can be estimated by measuring the output state of the channel,  $\rho_\theta$ .

the channel by applying it to a quantum state  $\rho_{\text{in}}$  to obtain the output state  $\rho_\theta = \mathcal{C}_\theta(\rho_{\text{in}})$  that hence acquires a dependence on the parameter  $\theta$ . By measuring this output state, information about the parameter can be obtained, allowing one to estimate the parameter.

The uncertainty, quantified as the root-mean-square error,  $\delta\hat{\theta}$  of an unbiased estimator  $\hat{\theta}$  of  $\theta$  can be bounded by the quantum Cramér-Rao bound [35, 36] as

$$\delta\hat{\theta} \geq \frac{1}{\sqrt{J_\theta(\rho_\theta)}}, \quad (1)$$

where  $J_\theta(\rho_\theta)$  is the quantum Fisher information (QFI) of the output state  $\rho_\theta$  with respect to the parameter  $\theta$ . For single-parameter estimation, which is the scope of this work, the quantum Cramér-Rao bound is known to be achievable [35, 36], making the QFI a central quantity in order to quantify the ability to estimate an unknown parameter. The QFI  $J_\theta(\rho_\theta)$  of the output state  $\rho_\theta$  (which is assumed to be continuously differentiable) can be written as [37]

$$J_\theta(\rho_\theta) := 2 \sum_{\substack{i,j: \\ \lambda_{\theta,i} + \lambda_{\theta,j} > 0}} \frac{|\langle \lambda_{\theta,j} | \partial_\theta \rho_\theta | \lambda_{\theta,i} \rangle|^2}{\lambda_{\theta,i} + \lambda_{\theta,j}}, \quad (2)$$

where the  $\lambda_{\theta,i}$ 's and  $|\lambda_{\theta,i}\rangle$ 's are the eigenvalues and corresponding (normalized) eigenvectors of  $\rho_\theta$ , so that  $\rho_\theta = \sum_i \lambda_{\theta,i} |\lambda_{\theta,i}\rangle \langle \lambda_{\theta,i}|$ . In this work we shall use another expression of the QFI introduced in Ref. [38],

$$J_\theta(\rho_\theta) = 4 \min_{\{|\psi_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \sum_{j=1}^q |\dot{\psi}_{\theta,j}\rangle \langle \dot{\psi}_{\theta,j}| \right), \quad (3)$$

where  $q$  can be any fixed value no smaller than the rank of  $\rho_\theta$ , and where the minimization is over all sets  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  of  $q$  (potentially unnormalized, not necessarily orthogonal) vectors that form a so-called ensemble decomposition of size  $q$  of  $\rho_\theta$  [38], such that  $\rho_\theta = \sum_{j=1}^q |\psi_{\theta,j}\rangle \langle \psi_{\theta,j}|$ , and where all  $|\psi_{\theta,j}\rangle$ 's are themselves continuously differentiable. The dot denotes the derivation with respect to  $\theta$ ,  $|\dot{\psi}_{\theta,j}\rangle = \partial_\theta |\psi_{\theta,j}\rangle$ .

In the simple scenario of Fig. 1, the channel  $\mathcal{C}_\theta$  is probed only once and it is only possible to play with the initial probe state  $\rho_{\text{in}}$  to maximize the QFI of the output state, and thus the precision of estimation. However, as soon as we consider scenarios in which  $\mathcal{C}_\theta$  can be used, or probed, several times, more possibilities arise. If one independently probes  $N$  copies of the channel  $\mathcal{C}_\theta$  one obtains the output state  $\rho_\theta^{\otimes N}$ , for which  $J_\theta(\rho_\theta^{\otimes N}) = NJ_\theta(\rho_\theta)$ , so that the Cramér-Rao bound gives

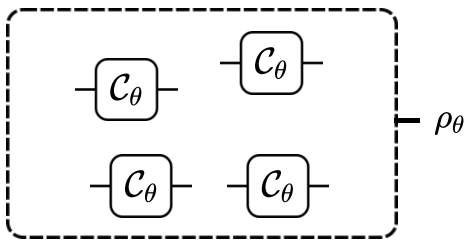


Figure 2. Abstract representation of a strategy that connects  $N = 4$  instances of the quantum channel  $\mathcal{C}_\theta$  in order to produce the output state  $\rho_\theta$ . In this paper we study the precision of estimating  $\theta$  from  $\rho_\theta$  given access to different classes of such strategies.

$\delta\hat{\theta} \geq \frac{1}{\sqrt{N J_\theta(\rho_\theta)}}$ . In addition to playing with the initial probe state, however, one can also change the structure of the probing strategy in order to maximize the QFI of the output state. For instance, it is well-known that, by entangling several probes, one can increase the precision of an estimate beyond the classical square-root scaling in  $N$  [39], even achieving Heisenberg scaling—a quadratic advantage in scaling. Compared to the strategy shown in Fig. 1, it can likewise be advantageous to consider probes that are entangled with additional ancillary systems that the channel is not applied to [40, 41]. More generally, given  $N$  uses of, or “queries” to, the quantum channel  $\mathcal{C}_\theta$ , one may wonder what the best circuit architecture, or strategy, is that maximizes the precision of estimation of  $\theta$ ; this scenario is depicted abstractly in Fig. 2.

This question and the underlying metrological task, which we will formalize further in Sec. IV, is the central object of study in this paper. Thus far, most approaches to quantum metrology have considered the most general strategy to be described by a standard, causally ordered quantum circuit. However, one can also envisage—as we will do here—more general compositional strategies using indefinite causal order, and study how useful this new resource is for improving the precision of parameter estimation.

## B. The process matrix formalism

In order to describe more general metrological strategies, we will make use of the process matrix formalism that was introduced in Ref. [1] to describe processes with potentially indefinite causal order. This formalism generalizes quantum theory by relaxing the assumption of a well-defined global causal order between different operations. Standard, causally definite quantum theory is instead only assumed to hold locally: operations are assigned to local closed and spatially separated laboratories described by quantum laws, but no assumption of a predefined global causal structure is made.

In the general scenario described by the process matrix formalism,  $N$  quantum operations  $\mathcal{A}_k$  (with  $1 \leq k \leq N$ ) are considered. Each operation is performed once and only once on the physical system received in its input Hilbert space  $\mathcal{H}^{A_k^I}$ , and the resulting system is sent out in the output Hilbert

space  $\mathcal{H}^{A_k^O}$ .<sup>1</sup> The operations are described by completely positive (CP) maps  $\mathcal{A}_k : \mathcal{L}(\mathcal{H}^{A_k^I}) \rightarrow \mathcal{L}(\mathcal{H}^{A_k^O})$ , where  $\mathcal{L}(\mathcal{H}^X)$  denotes the space of linear operators on  $\mathcal{H}^X$ . The overall process describes how the  $N$  operations are related, be it in a definite or indefinite causal order, and outputs a system in the global causal future of all the operations, described by some Hilbert space  $\mathcal{H}^F$ .<sup>2</sup> This process can equivalently be seen as a higher-order map, or supermap, that maps the  $N$  operations  $\mathcal{A}_k$  to the state output in the global future. By considering its Choi–Jamiołkowski representation [43], the process can be uniquely described by a matrix  $W \in \mathcal{L}(\mathcal{H}^{A_N^I O^F})$ , its *process matrix*, where we use the shorthand notation  $\mathcal{H}^{XY} := \mathcal{H}^X \otimes \mathcal{H}^Y$ ,  $\mathcal{H}^{A_k^I O} := \mathcal{H}^{A_k^I A_k^O}$ , and  $\mathcal{H}^{A_K^I O} := \bigotimes_{k \in \mathcal{K}} \mathcal{H}^{A_k^I O}$ , for any  $\mathcal{K} \subseteq \mathcal{N} := \{1, \dots, N\}$ . The fact that the overall process must yield a proper, well-defined supermap, but without making any assumption of a predefined global causal order between the operations, imposes some constraints on the process matrix  $W$  that must be satisfied for it to be valid; we will detail these in Sec. V. Within the set of all valid process matrices one can identify specific relevant classes of process matrices (for instance, those compatible with a definite causal order) obtained by imposing further constraints on  $W$ , which will also be detailed in Sec. V. We shall denote by  $\mathbf{W}$  a generic class of processes.

When considering the scenario of Fig. 2, we are typically interested in expressing the output state  $\rho_\theta$  as a function of the channels  $\mathcal{C}_\theta$  and the process matrix  $W$  describing a given strategy. To this end, let us first introduce the Choi matrix [43]  $C_\theta$  of the channel  $\mathcal{C}_\theta : \mathcal{L}(\mathcal{H}^{A_k^I}) \rightarrow \mathcal{L}(\mathcal{H}^{A_k^O})$ , that connects an input space  $\mathcal{H}^{A_k^I}$  to an output space  $\mathcal{H}^{A_k^O}$ , as

$$C_\theta := \sum_{i,j} |i\rangle\langle j|^{A_k^I} \otimes \mathcal{C}_\theta(|i\rangle\langle j|^{A_k^I}) \in \mathcal{L}(\mathcal{H}^{A_k^I O}), \quad (4)$$

where  $\{|i\rangle^X\}_i$  denotes a fixed orthonormal basis, which we refer to as the computational basis of  $\mathcal{H}^X$ , and where we use the short-hand notation  $|i\rangle\langle j|^X := |i\rangle^X \langle j|^X$ . Throughout this paper we will refer to the channel  $\mathcal{C}_\theta$  and its Choi matrix  $C_\theta$  indifferently, as Eq. (4) describes an isomorphism that can be inverted to express  $\mathcal{C}_\theta$  in terms of  $C_\theta$ .

When using  $N$  copies of the channel, as in Fig. 2, one can then rewrite the output state  $\rho_\theta$  using the so-called link product

<sup>1</sup> All Hilbert spaces are assumed here to be finite-dimensional. For simplicity we shall typically consider that the operations input and output Hilbert spaces are of the same dimension,  $d$ .

<sup>2</sup> Although the original formulation of the process matrix framework did not involve the output system in the space  $\mathcal{H}^F$ , it is rather easy to extend the formalism so as to include it [42]. It is also possible to introduce some space  $\mathcal{H}^P$  in which some input system could be prepared, in the global causal past of all other operations. As in Refs. [15, 19], however, we do not define such a global past, since any fixed initial state (as one has in the metrological scenario we consider) can be incorporated directly into the process itself.

Note also that in general the operations  $\mathcal{A}_k$  can be extended to also act on auxiliary spaces  $\mathcal{H}^{A_k^{I' O'}}$ , i.e., the supermap can be applied to CP maps  $\mathcal{A}_k : \mathcal{L}(\mathcal{H}^{A_k^I A_k^{I'}}) \rightarrow \mathcal{L}(\mathcal{H}^{A_k^O A_k^{O'}})$ .

as

$$\rho_\theta = C_\theta^{\otimes N} * W. \quad (5)$$

The link product between two operators  $A \in \mathcal{L}(\mathcal{H}^{XY})$  and  $B \in \mathcal{L}(\mathcal{H}^{YZ})$  that share a common system  $Y$  associated to  $\mathcal{H}^Y$  is defined as [44, 45]

$$A * B := \text{Tr}_Y[(A^{T_Y} \otimes \mathbb{1}^Z)(\mathbb{1}^X \otimes B)] \in \mathcal{L}(\mathcal{H}^{XZ}), \quad (6)$$

where  $\text{Tr}_Y$  and  $T_Y$  denote, respectively, the partial trace and partial transpose (in the computational basis) with respect to the Hilbert space  $\mathcal{H}^Y$ , and  $\mathbb{1}$  is the identity operator in the appropriate space (indicated here by the superscripts).

### C. Example: the quantum switch

#### 1. Description of the dynamics

The quantum switch [2] is the canonical example of a process with indefinite causal order, and can be described as follows. Two operations  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are applied locally to a target system initially in the state  $|\psi_{\text{in}}\rangle^t \in \mathcal{H}^t$  in an order that depends on a control qubit in the space  $\mathcal{H}^c$ , after which both the target system and the control qubit are sent to a region corresponding to the global future  $F$  with Hilbert space  $\mathcal{H}^F := \mathcal{H}^{F^c} \otimes \mathcal{H}^{F^t}$ . If the control qubit is in the state  $|0\rangle^c$ , then  $\mathcal{A}_1$  is applied to the target system before  $\mathcal{A}_2$ , while if the control qubit is in the state  $|1\rangle^c$ , then  $\mathcal{A}_2$  is applied before  $\mathcal{A}_1$ . By initializing the control qubit in the superposed state  $|+\rangle^c = \frac{1}{\sqrt{2}}(|0\rangle^c + |1\rangle^c)$ , and in the case where  $\mathcal{A}_i$  (for  $i = 1, 2$ ) is a unitary operation  $\rho \mapsto U_{A_i} \rho U_{A_i}^\dagger$ , with  $U_{A_i}$  a unitary, the dynamics of the quantum switch is given by

$$\begin{aligned} & |+\rangle^c \otimes |\psi_{\text{in}}\rangle^t \\ & \rightarrow \frac{1}{\sqrt{2}}(|0\rangle^c \otimes U_{A_2} U_{A_1} |\psi_{\text{in}}\rangle^t + |1\rangle^c \otimes U_{A_1} U_{A_2} |\psi_{\text{in}}\rangle^t). \end{aligned} \quad (7)$$

The switch can thus be interpreted as a coherent control of the two possible causal orders between  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . Although its expression is slightly more involved, the dynamics can also be written for nonunitary operations  $\mathcal{A}_i$ , with a similar interpretation [2].

#### 2. Process matrix description

The process matrix for the quantum switch process can be written as [8]

$$\begin{aligned} W^{\text{QS}} &= |w\rangle\langle w| \\ \text{with } |w\rangle &= \frac{1}{\sqrt{2}}(|\psi_{\text{in}}\rangle^{A_1^I} |\mathbb{1}\rangle^{A_1^O A_2^I} |\mathbb{1}\rangle^{A_2^O F^t} |0\rangle^{F^c} \\ & \quad + |\psi_{\text{in}}\rangle^{A_2^I} |\mathbb{1}\rangle^{A_2^O A_1^I} |\mathbb{1}\rangle^{A_1^O F^t} |1\rangle^{F^c}), \end{aligned} \quad (8)$$

with the tensor products left implicit for readability, and where  $|\mathbb{1}\rangle^{XY} := \sum_i |i\rangle^X \otimes |i\rangle^Y$  is defined for two isomorphic spaces

$\mathcal{H}^X$  and  $\mathcal{H}^Y$ , whose computational bases  $\{|i\rangle^X\}_i$  and  $\{|i\rangle^Y\}_i$  are in one-to-one correspondence. The terms of the form  $|\mathbb{1}\rangle^{XY}$  can be understood as representing an identity channel between  $\mathcal{H}^X$  and  $\mathcal{H}^Y$ . Eq. (8) involves the coherent superposition of two terms that correspond to the two possible orders in which  $\mathcal{A}_1$  comes before  $\mathcal{A}_2$  and in which  $\mathcal{A}_2$  comes before  $\mathcal{A}_1$ , in agreement with Eq. (7).

The quantum switch has been shown to provide an advantage over processes with definite causal order for different tasks in several subfields of quantum information, including quantum computation [5, 46] and quantum communication [6, 9]. Recently, several attempts to demonstrate advantages in quantum metrology using the quantum switch have been presented. Some of these fit within the standard framework of quantum metrology we described above [11, 12, 14, 16–18, 23], while others have considered more particular scenarios involving, for example, products of position and momentum displacement operators [13]. These works, particularly the former ones in the more standard quantum metrology scenario, motivate our work, and we will review some of them further in what follows.

## III. RE-EXAMINING THE LITERATURE

Here we examine more carefully some of the supposed advantages in quantum metrology obtained from indefinite causal order. Refs. [11, 12] exemplify the attempts to use the quantum switch for metrology in the standard metrological setting, and by studying these examples we will argue that it is misleading to claim that indefinite causal order would be a resource for quantum metrology with respect to the evidence they provide.

### A. Review of some previous work

In Ref. [12] the author considers the  $d$ -dimensional depolarizing channel  $\mathcal{D}_\theta$  defined as

$$\mathcal{D}_\theta(\rho) = \theta\rho + (1 - \theta)\frac{\mathbb{1}}{d}, \quad (9)$$

for any (normalized)  $d$ -dimensional density matrix  $\rho$ , and where  $\theta \in [0, 1]$  parametrizes the channel. They study the following metrological comparison: given two instances of  $\mathcal{D}_\theta$ , compare the QFI of the output state produced by the quantum switch,<sup>3</sup>  $J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}})$ , and that produced by the sequential quantum circuit that composes the instances of  $\mathcal{D}_\theta$  one after the other (described by the process matrix  $W^{\text{Seq}}$ ; see Appendix A 1),  $J_\theta(D_\theta^{\otimes 2} * W^{\text{Seq}})$ . On the grounds that  $J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}}) > J_\theta(D_\theta^{\otimes 2} * W^{\text{Seq}})$  for all  $\theta < 1$ , the author concludes that the quantum switch is an aid for channel probing, in the sense that causal indefiniteness leads to higher QFI,

<sup>3</sup> Both Refs. [11, 12] actually consider a quantum switch with a control qubit initialized in a more general state of the form  $\sqrt{p}|0\rangle^c + \sqrt{1-p}|1\rangle^c$ . The claimed advantages are found to be the largest for  $p = 1/2$ ; for simplicity we restrict here the discussion to that case.



and hence better performance. For the comparison of these two specific processes, the quantum switch  $W^{\text{QS}}$  and the sequential circuit  $W^{\text{Seq}}$ , this statement is strictly true, but insofar as it concerns causal indefiniteness more generally it appears rather limited or even misleading. One may wonder, for instance, whether the quantum switch outperforms all strategies with a definite causal order—as one would like if one is to claim indefinite causal order is a resource for channel probing—or simply the specific sequential strategy considered.

The same kind of comparison was made in Ref. [11], where this time the quantum switch and the sequential strategy are compared in a metrological task related to quantum thermometry. The same scenario as in Ref. [12] is considered, with the depolarizing channel  $\mathcal{D}_\theta$  replaced by a qubit thermalizing channel  $\mathcal{T}_\theta$ , where  $\theta > 0$  is here the temperature of the bath the qubit is probing. The channel  $\mathcal{T}_\theta$  is defined by the following Kraus operators:

$$\begin{aligned} K_{\theta,0} &= \sqrt{p_\theta} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix}, \quad K_{\theta,1} = \sqrt{p_\theta} \begin{pmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{pmatrix}, \\ K_{\theta,2} &= \sqrt{1-p_\theta} \begin{pmatrix} \sqrt{1-\lambda} & 0 \\ 0 & 1 \end{pmatrix}, \quad K_{\theta,3} = \sqrt{1-p_\theta} \begin{pmatrix} 0 & 0 \\ \sqrt{\lambda} & 0 \end{pmatrix}, \end{aligned} \quad (10)$$

so that  $\mathcal{T}_\theta(\rho) = \sum_{i=0}^3 K_{\theta,i} \rho K_{\theta,i}^\dagger$ , where  $p_\theta = 1/(1+e^{-\epsilon/\theta})$ ,  $\lambda = 1 - e^{-t/\tau}$ , with  $\epsilon > 0$  the energy gap of the probe,  $t > 0$  the interaction time between the probe and the bath, and  $\tau > 0$  the relaxation time of the bath. In the limit  $t \rightarrow \infty$ , for which  $\lambda \rightarrow 1$ , any input state is thermalized so that the thermalizing channel becomes constant:  $\mathcal{T}_\theta(\rho) = p_\theta |0\rangle\langle 0| + (1-p_\theta) |1\rangle\langle 1|$ ,  $\forall \rho$  (with  $\text{Tr} \rho = 1$ ). In this limit, by showing that  $J_\theta(\mathcal{T}_\theta^{\otimes 2} * W^{\text{QS}}) > J_\theta(\mathcal{T}_\theta^{\otimes 2} * W^{\text{Seq}})$  for all  $\theta$  (and for the input state of the quantum switch  $|\psi_{\text{in}}\rangle = |0\rangle$ , see Appendix A 2), the authors conclude again that indefinite causal order is a resource in quantum thermometry. As before, the scope of this comparison is rather limited by considering only a single causally-definite strategy for comparison with the quantum switch.

### B. Comment and critique

As shown in Appendix A, for each of the two tasks introduced in [11, 12], we could come up with an example of a process with definite causal order that outperforms the quantum switch, yielding a larger QFI. The two corresponding circuits are depicted in Fig. 3. Both are parallel circuits, which ensures they are compatible with a definite causal order. The circuit presented in Fig. 3a consists of two independent uses of a sub-circuit, where the probe system sent through the depolarizing channel is initially prepared along with an auxiliary system, in a maximally entangled state  $|\phi^+\rangle := \frac{1}{\sqrt{d}} |\mathbb{1}\rangle\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle \otimes |i\rangle$ . We term this sub-circuit the ‘‘Choi circuit’’, since its output state is proportional to the Choi matrix of the channel  $\mathcal{D}_\theta$ . The overall circuit is thus termed the 2-Choi circuit, because it consists of two independent uses of the Choi circuit. The circuit presented in Fig. 3b simply applies the two copies of

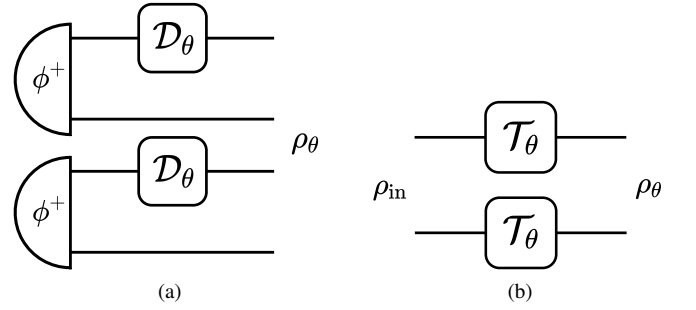


Figure 3. (a): Optimal strategy for the parameter estimation of  $N = 2$  depolarizing channels as considered in [12]. This so-called ‘‘2-Choi circuit’’ is made of a parallel use of the two depolarizing channels represented by the rounded square boxes, with maximally entangled input states  $|\phi^+\rangle := \frac{1}{\sqrt{d}} |\mathbb{1}\rangle\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle \otimes |i\rangle$ . (b): Optimal strategy for  $N = 2$  thermalizing channels as defined in [11] (in the limit  $t \rightarrow \infty$ ). This circuit simply consists of a parallel use of the two thermalizing channels represented by the rounded square boxes.

the thermalizing channel in parallel, with no constraint on the input state of the probe system.

These two examples show that a fairer comparison should be made in order to claim an advantage of the quantum switch, or more generally of causally indefinite order processes, in the corresponding tasks. To claim an advantage of the quantum switch in such a scenario, one should instead make a comparison between its performance and the performance of the whole class of processes with a definite causal order. As proposed in [13], one way to do this is to derive a bound on the QFI of any process with a definite causal order, and then show that the bound can be outperformed by the quantum switch. In general, however, analytically finding such a bound is a difficult task. Another approach, based on semidefinite programming (SDP) techniques (which are commonly used in the study of causally indefinite processes), is to optimize the QFI of the output state over classes of strategies with definite or indefinite causal orders. This was first proposed in the framework introduced in [15] for strategies with a definite causal order, and then extended to some different classes of strategies with indefinite causal order in [19].

In what follows, we give a formal definition of the metrological task we consider and then show how strategies can be compared in this task using SDPs.

## IV. COMPARING THE METROLOGICAL PERFORMANCE OF WHOLE CLASSES OF PROCESSES

As motivated above, a better understanding of the possible advantage of indefinite causal order processes in quantum metrology requires a broader approach that allows for comparison of classes of processes with definite causal order with classes of processes with indefinite causal order. This is the purpose of the framework introduced in [15, 19]. We first review this work and then extend it, considering recently introduced classes of circuits, in order to better qualify the advantage of indefinite causal order processes in quantum metrology.

### A. The metrological task: optimizing the quantum Fisher information over a full class of strategies

As already introduced, given  $N$  queries to a quantum channel  $C_\theta$  that depends on an unknown parameter  $\theta$ , we want to find the strategy that maximizes the QFI of its output state  $\rho_\theta$ , over a given class of strategies  $\mathbf{W}$ . More formally, a strategy can be described as a process matrix  $W \in \mathcal{L}(\mathcal{H}^{A_N^I O F})$ , that can be fed  $N$  copies of  $C_\theta$  (in the Choi representation)—i.e., with the  $(d^{2N} \times d^{2N})$  Choi matrix  $C_\theta^{\otimes N} \in \mathcal{L}(\mathcal{H}^{A_N^I O F})$ —in order to produce an output state  $\rho_\theta = C_\theta^{\otimes N} * W \in \mathcal{L}(\mathcal{H}^F)$ . Mathematically, this boils down to determining

$$J_\theta^{\mathbf{W}}(C_\theta^{\otimes N}) := \max_{W \in \mathbf{W}} J_\theta(C_\theta^{\otimes N} * W), \quad (11)$$

where  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$  is the optimal QFI over the class of strategies  $\mathbf{W}$  using  $N$  queries to  $C_\theta$ , with respect to the parameter  $\theta$ . Following [15, 19], it is possible to rewrite  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$  as an SDP for several relevant classes  $\mathbf{W}$ , allowing efficient numerical methods to be used to calculate  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$ .

### B. Solving the task through SDP optimization

The framework introduced in [15, 19] proposes a method to rewrite Eq. (11) as an SDP, which can then be solved numerically to obtain an optimal strategy for the metrological task defined above, over a given class of strategies with or without a definite causal order. In order to generalize this framework to new classes of strategies  $\mathbf{W}$ , we first give a sketch of the arguments which will next serve as a base for the extension to the new classes. For further details, some clarifications, and proofs of the arguments, see Appendix B 1.

First, we use the convexity of the QFI to restrict the optimization over the strategies  $W \in \mathbf{W}$  in Eq. (11) to pure (i.e., rank-one) process matrices, such that  $W = |w\rangle\langle w| \in \mathbf{W}$ ; see Appendix B 1 a. For all the classes considered in [15, 19], as well as for the QC-QCs that we shall introduce below, this step can be done without loss of generality since any process matrix in these classes can be purified within the same class, up to enlarging the future space  $\mathcal{H}^F$  (which our definitions of the different classes allow for; see below). But, in general, this is not a trivial step, as it might be that the purifications of some process matrices in a given class of strategies belong to a larger class of strategies. This is in particular the case for the class of QC-CCs presented below.<sup>4</sup> The approach used here then only allows one to obtain the optimal QFI within this larger purifying class—which thus only gives an upper bound for the QFI with respect to the class initially considered.

Restricting to pure process matrices allows one, in the expression of the QFI of the output state  $\rho_\theta = C_\theta^{\otimes N} * W =$

$C_\theta^{\otimes N} * |w\rangle\langle w|$  from Eq. (3), to turn the optimization over ensemble decompositions of size  $q$  of  $\rho_\theta$ ,  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$ , into an optimization over ensemble decompositions of the same size  $q$ ,  $\{|C_{\theta,j}\rangle\}_{j=1}^q$ , of the channels' operator  $C_\theta^{\otimes N}$  directly:

$$\begin{aligned} J_\theta(\rho_\theta) &= 4 \min_{\{|\psi_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \sum_{j=1}^q |\psi_{\theta,j}\rangle\langle\psi_{\theta,j}| \right) \\ &= 4 \min_{\{|C_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \left( \sum_{j=1}^q |\dot{C}_{\theta,j}\rangle\langle\dot{C}_{\theta,j}| \right) * W \right). \end{aligned} \quad (12)$$

See Appendix B 1 b for more details. Here we assume that the dependency of the channels on  $\theta$  is smooth enough, so that similarly to  $\min_{\{|\psi_{\theta,j}\rangle\}}$ , the minimum  $\min_{\{|C_{\theta,j}\rangle\}}$  is taken over continuously differentiable decompositions.

Then, any such ensemble decomposition of size  $q$  of  $C_\theta^{\otimes N}$  can be expressed in terms of a fixed ensemble decomposition of size  $q$  of  $C_\theta^{\otimes N}$ ,  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$ , and a (sufficiently smooth)  $q \times q$  unitary  $V_\theta$  that relates the two decompositions. This allows us to replace the optimization over  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  by an optimization over  $V_\theta$ ; see Appendix B 1 c. By realizing that the objective function of the optimization only depends on the  $q \times q$  Hermitian matrix  $h = iV_\theta V_\theta^\dagger$ , the optimization over  $V_\theta$  is, in turn, replaced by an optimization over such Hermitian matrices  $h$ , cf. Appendix B 1 d. One thereby obtains the expression

$$J_\theta(\rho_\theta) = \min_{h \in \mathbb{H}_q} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right), \quad (13)$$

where  $\mathbb{H}_q$  denotes the set of  $q \times q$  Hermitian matrices, and  $\tilde{W} = \text{Tr}_F W$ . Here we introduced the so-called ‘‘performance operator’’ [15, 47]

$$\Omega_\theta(h) = 4 \left( (\dot{C}_\theta^0 - iC_\theta^0 h) (\dot{C}_\theta^0 - iC_\theta^0 h)^\dagger \right)^T, \quad (14)$$

where  $T$  denotes transposition in the computational basis, and with  $C_\theta^0 = \left( |C_{\theta,1}^0\rangle, \dots, |C_{\theta,q}^0\rangle \right)$  being the  $d^{2N} \times q$  matrices whose column vectors are the  $(d^{2N}$ -dimensional) elements of the fixed decomposition  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  of  $C_\theta^{\otimes N}$  (and with  $\dot{C}_\theta^0$  its derivative with respect to  $\theta$ ). As one can observe, the QFI of  $\rho_\theta$  is obtained by optimizing over  $\tilde{W}$ , the process matrix of the strategy with the global future traced out. Thus, the action of  $W$  on the global future  $F$ , and in particular the dimension of the Hilbert space  $\mathcal{H}^F$ , do not have to be specified at this level. One only needs to fix the dimension of  $\mathcal{H}^F$  when one wants to purify  $\tilde{W}$  to obtain the underlying strategy in full.

Then, injecting Eq. (13) into Eq. (11), one obtains the expression of the optimal QFI over a given class of strategies  $\mathbf{W}$ :

$$J_\theta^{\mathbf{W}}(C_\theta^{\otimes N}) = \max_{\tilde{W} \in \text{Tr}_F \mathbf{W}} \min_{h \in \mathbb{H}_q} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right), \quad (15)$$

with  $\text{Tr}_F \mathbf{W} := \{\text{Tr}_F W, W \in \mathbf{W}\}$ . Finally, because the objective function is convex in  $h$  and linear (thus concave)

<sup>4</sup> For instance, the ‘‘classical switch’’ (the classical counterpart of the quantum switch) is a QC-CC whose purification is a QC-QC but no longer a QC-CC [34].

in  $\tilde{W}$ , and assuming that  $\mathbf{W}$  is compact (as will be the case for all classes under consideration here), one can exchange the minimization with the maximization using Fan's minimax theorem [48], and obtain (see Appendix B 1 e)

$$J_{\theta}^{\mathbf{W}}(C_{\theta}^{\otimes N}) = \min_{h \in \mathbb{H}_{\theta}} \max_{\tilde{W} \in \text{Tr}_F \mathbf{W}} \text{Tr} \left( \Omega_{\theta}(h) \tilde{W} \right). \quad (16)$$

Since this equation is a min max problem, it is not straightforward in general to solve it numerically. As we shall see, the classes  $\mathbf{W}$  that we will be interested in can all be characterized with semidefinite constraints, making the inner maximization problem an SDP. By first fixing  $h$  and using duality to turn the maximization into a minimization, and after conveniently linearizing the (*a priori* quadratic in  $h$ ) constraints that involve  $\Omega_{\theta}(h)$ , one can then rewrite Eq. (16) itself as an SDP (Appendix B 1 f). Solving this SDP gives both  $J_{\theta}^{\mathbf{W}}(C_{\theta}^{\otimes N})$ , the QFI optimized over the class  $\mathbf{W}$ , and  $h^{(\text{opt})}$ , the value of the optimal  $h$  (Appendix B 1 g). Fixing  $h = h^{(\text{opt})}$  and imposing an extra condition on  $\tilde{W}$  that we discuss in Appendix B 1 h (see also Appendix B 3), one can then also recover an optimal strategy  $\tilde{W}^{(\text{opt})}$  from Eq. (15). An optimal strategy  $W^{(\text{opt})}$  can then finally be obtained by purifying  $\tilde{W}^{(\text{opt})}$  using the global future  $F$ . As mentioned previously, since not all QC-CCs are purifiable as QC-CCs this approach computes the optimal QFI for a larger class than QC-CC and that includes all the purifications of QC-CCs. We denote the QFI of this larger class by  $J_{\theta, \text{purif}}^{\text{QC-CC}}(C_{\theta}^{\otimes N})$ , and emphasize that it is an upper bound of  $J_{\theta}^{\text{QC-CC}}(C_{\theta}^{\otimes N})$ .

## V. CLASSES OF STRATEGIES

With the metrological task defined and the methodology to solve it presented, we review in this section the different classes of strategies  $\mathbf{W}$  that we are going to compare and their characterizations in terms of SDP constraints. Further details and, in particular, explicit formulations of the (primal and dual) SDP problems computing  $J_{\theta}^{\mathbf{W}}(C_{\theta}^{\otimes N})$  for the different classes can be found in Appendix B 2.

As already noted, in the calculation of  $J_{\theta}^{\mathbf{W}}(C_{\theta}^{\otimes N})$  one only needs characterizations of the sets  $\text{Tr}_F \mathbf{W}$ .<sup>5</sup> In the definitions of the classes  $\mathbf{W} \subseteq \mathcal{L}(\mathcal{H}^{A_N^{IO} F})$  that follow, one should thus understand the dimension of  $\mathcal{H}^F$  to be free, in contrast to that of the  $\mathcal{H}^{A_k^{IO}}$ , which are fixed by the dimension of  $C_{\theta}$ . That is, a process  $W \in \mathbf{W}$  (for each  $\mathbf{W}$  specified below) if, for some dimension of  $\mathcal{H}^F$ ,  $W$  satisfies the corresponding constraints we specify for  $\mathbf{W}$ .

### A. QC-Par, (conv)QC-FO, QC-Sup and Gen

One of the simplest classes of quantum strategies is to compose the  $N$  operations  $\mathcal{A}_k$  (with  $1 \leq k \leq N$ ) in parallel. In such a quantum circuit with operation used in parallel (QC-Par), each  $\mathcal{A}_k$  is applied to part of a multipartite state that is in general entangled across the  $\mathcal{A}_k$  and, eventually, with an ancillary system. A quantum channel can then map the resulting output state to the global future  $F$ . This class has been extensively studied and the corresponding class QC-Par of process matrices can be characterized as [45]

$$\text{QC-Par} = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_N^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} \text{Tr}_F W = W_{(I)} \otimes \mathbb{1}^{A_N^O}, \\ \text{Tr} W_{(I)} = 1, \\ W_{(I)} \geq 0 \end{array} \right. \end{array} \right\}, \quad (17)$$

where  $W_{(I)} \in \mathcal{L}(\mathcal{H}^{A_N^I})$  is simply an  $N$ -partite density matrix.

Another extensively studied class of strategies is quantum circuits with a fixed order (QC-FO) [34], also known as “quantum strategies” [49] or “quantum combs” [44]. This class corresponds to “standard” quantum circuits, where the  $\mathcal{A}_k$  are applied sequentially, potentially with intermediate quantum channels applied between the  $\mathcal{A}_k$  jointly on the outcome of the previous operation and some quantum memory. Such QC-FOs include parallel circuits as a special case, and their process matrices can be formally characterized as

$$\text{QC-FO} = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_N^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} \text{Tr}_F W = W_{(N)} \otimes \mathbb{1}^{A_N^O}, \\ \forall n = 1, \dots, N-1, \\ \quad \text{Tr}_{A_{n+1}^I} W_{(n+1)} = W_{(n)} \otimes \mathbb{1}^{A_n^O}, \\ \text{Tr} W_{(1)} = 1, \\ \forall n = 1, \dots, N, W_{(n)} \geq 0 \end{array} \right. \end{array} \right\}, \quad (18)$$

where  $W_{(n)} \in \mathcal{L}(\mathcal{H}^{A_{\{1, \dots, n-1\}}^I A_n^I})$ , for  $n = 1, \dots, N$ . Strictly speaking, the definition of QC-FO as just stated only includes circuits with the specific fixed causal order  $\mathcal{A}_1, \dots, \mathcal{A}_N$ . However, any other permutation of this order would still define a class of circuits with a fixed causal order. Thus, we define the classes QC-FO $_{\pi}$ , which are defined analogously to QC-FO, but which correspond to the causal orders  $\mathcal{A}_{\pi(1)}, \dots, \mathcal{A}_{\pi(N)}$ , with  $\pi$  a permutation of  $\mathcal{N}$ . Note that for each  $\pi$ , QC-FO $_{\pi}$  is convex; however, their union is not. Therefore we also define convQC-FO, the convex hull of all the quantum circuits with a fixed causal order, as

$$\text{convQC-FO} := \text{conv}(\cup_{\pi} \text{QC-FO}_{\pi}) = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_N^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} W = \sum_{\pi} q_{\pi} W_{\pi}, \\ \forall \pi, W_{\pi} \in \text{QC-FO}_{\pi}, \\ \sum_{\pi} q_{\pi} = 1, \forall \pi, q_{\pi} \geq 0 \end{array} \right. \end{array} \right\}, \quad (19)$$

<sup>5</sup> Note, in particular, that [19] made a simplifying assumption (see their Eq. (9)) about the structure of  $\text{Tr}_F \mathbf{W}$  for the classes they considered, but which is not satisfied by some of the classes we introduce here. We hence rederive the SDP problems for  $J_{\theta}^{\mathbf{W}}(C_{\theta}^{\otimes N})$  directly from the full characterizations of each  $\mathbf{W}$  we give here; cf. Appendix B 2.



where the sums are over all possible permutations  $\pi$  of  $\mathcal{N}$ . By definition, this class only comprises processes with a definite causal order, since they are defined as classical mixtures of processes with a fixed causal order.

In analogy with the quantum switch, one can construct processes that use superpositions of different fixed causal orders by using a control system, which is given to the global future  $F$ , to coherently control the order. Such strategies correspond to the class of quantum circuits with causal superposition (QC-Sup) that was introduced in [19], and which can be characterized as

$$\text{QC-Sup} = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_{\mathcal{N}}^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} \text{Tr}_F W = \sum_{\pi} q_{\pi} \tilde{W}_{\pi}, \\ \forall \pi, \tilde{W}_{\pi} \in \text{Tr}_F(\text{QC-FO}_{\pi}), \\ \sum_{\pi} q_{\pi} = 1, \forall \pi, q_{\pi} \geq 0 \end{array} \right. \end{array} \right\}. \quad (20)$$

Similarly to the characterization of convQC-FO, this characterization involves a convex mixture of processes with fixed causal order. Crucially, however, the convex decompositions apply to  $\text{Tr}_F W$  here, while for convQC-FO they applied to  $W$  directly. This explains the main difference between the two classes: QC-Sup allows for processes with causal superposition, and hence indefinite causal order, while convQC-FO does not.

One can also consider strategies corresponding to arbitrary process matrices, which are the most general way to consistently combine the operations  $\mathcal{A}_k$ . We will denote this class, which includes all strategies with indefinite causal order, as Gen, and it can be characterized as [8]

$$\text{Gen} = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_{\mathcal{N}}^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} \forall \emptyset \subsetneq \mathcal{K} \subseteq \mathcal{N}, \\ \prod_{k \in \mathcal{K}} [1 - A_k^{IO}] (\text{Tr}_{A_{\mathcal{N} \setminus \mathcal{K}}^{IO} F} W) = 0, \\ \text{Tr} W = \prod_{k \in \mathcal{N}} d_{A_k^{IO}} \end{array} \right. \end{array} \right\}, \quad (21)$$

with the generic short-hand notation  $d_X := \dim(\mathcal{H}^X)$  and the ‘‘trace-out-and-replace’’ notation defined as

$${}_X W := (\text{Tr}_X W) \otimes \frac{\mathbb{1}^X}{d_X}, \quad [1-X] W := W - {}_X W. \quad (22)$$

(The second part of the previous definition can be used recursively, in a commutative manner: e.g.,  $[1-X][1-Y]W = [1-X]([1-Y]W) = [1-Y]([1-X]W) = W - {}_X W - {}_Y W - {}_{XY} W$ .) Note that not all strategies in Gen are known to have a physical representation in terms of a generalized form of quantum circuit or otherwise.

## B. QC-CC and QC-QC

The classes of strategies introduced above were all considered in [19] and, with the exception of QC-Sup, were previously well studied. Here we present two further classes that

were recently introduced in [34] and which are of particular physical interest.

The first of these are quantum circuits with classical control of causal order (QC-CC), which attempt to generalize convQC-FO as the most general strategies compatible with a well-defined causal order. Contrary to QC-FOs, the causal order of a QC-CC is not necessarily fixed *a priori*, but can be established on the fly: a classical control system is used at each step in the circuit to control which operation (amongst those not yet applied) to apply next. Such QC-CCs thus allow the causal order to be dynamically controlled. QC-CCs can be characterized in terms of SDP constraints involving (positive semidefinite) matrices  $W_{(k_1, \dots, k_n)} \in \mathcal{L}(\mathcal{H}^{A_{\{k_1, \dots, k_{n-1}\}}^{IO} A_{k_n}^I})$  and  $W_{(k_1, \dots, k_N, F)} \in \mathcal{L}(\mathcal{H}^{A_{\mathcal{N}}^{IO} F})$  as

$$\text{QC-CC} = \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}^{A_{\mathcal{N}}^{IO} F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} W = \sum_{(k_1, \dots, k_N)} W_{(k_1, \dots, k_N, F)}, \\ \forall (k_1, \dots, k_N), \\ \text{Tr}_F W_{(k_1, \dots, k_N, F)} = W_{(k_1, \dots, k_N)} \otimes \mathbb{1}^{A_{k_N}^{IO}}, \\ \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ \sum_{k_{n+1}} \text{Tr}_{A_{k_{n+1}}^I} W_{(k_1, \dots, k_n, k_{n+1})} \\ = W_{(k_1, \dots, k_n)} \otimes \mathbb{1}^{A_{k_n}^{IO}}, \\ \sum_{k_1} \text{Tr} W_{(k_1)} = 1, \\ \forall n = 1, \dots, N, \forall (k_1, \dots, k_n), \\ W_{(k_1, \dots, k_n)} \geq 0, W_{(k_1, \dots, k_N, F)} \geq 0 \end{array} \right. \end{array} \right\}, \quad (23)$$

where for all  $n$ -tuples  $(k_1, \dots, k_n)$ , we implicitly assume that  $k_i \neq k_j$  for  $i \neq j$ . For  $N \leq 3$ , it was shown [50] that QC-CC constitutes the class of all circuits with a definite causal order (so-called ‘‘causally separable’’ processes), while it is still an open question for  $N > 3$ .

The second new class we consider are the quantum circuits with quantum control of causal order (QC-QC), which can be thought of as the quantum counterpart of QC-CCs. QC-QCs use a quantum control system to coherently control which operation is applied at each step, allowing for both coherent causal superposition and dynamical control of causal order. QC-QCs may thus exhibit indefinite causal order, and are the most general class of processes known to have a clear physical interpretation [34, 51], which gives them particular practical importance for understanding any potential advantages from indefinite causal order in quantum metrology. This class can be characterized in terms of (positive semidefinite) matrices  $W_{(\mathcal{K}_{n-1}, k_n)} \in \mathcal{L}(\mathcal{H}^{A_{\mathcal{K}_{n-1}}^{IO} A_{k_n}^I})$ , for all strict subsets  $\mathcal{K}_{n-1}$  of

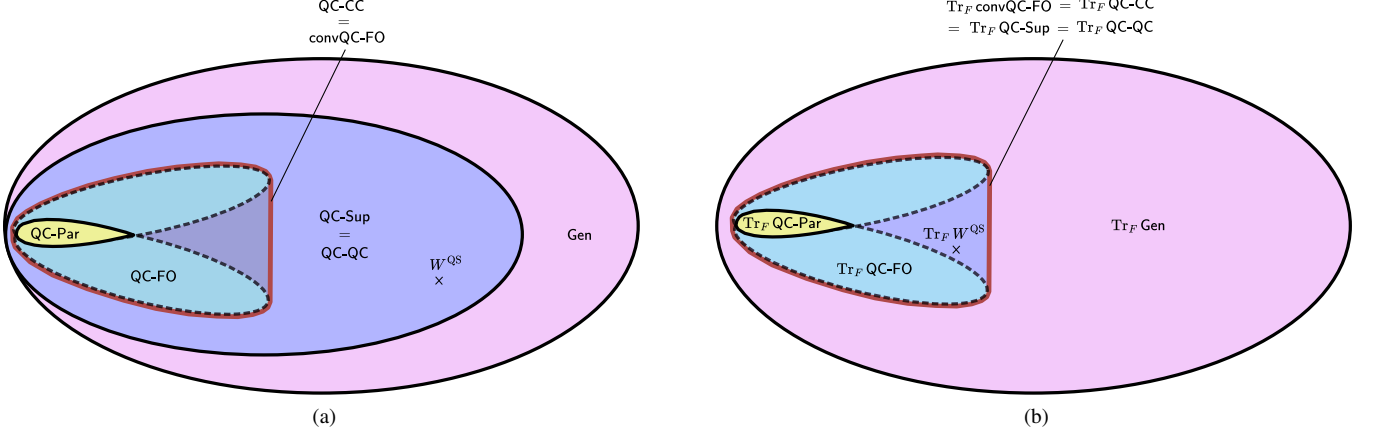


Figure 4. Relations between the different classes of strategies with and without definite causal order for  $N = 2$ , in the cases where the global future  $F$  is (a) not traced out; (b) traced out.

$\mathcal{N}$  and all  $k_n \in \mathcal{N} \setminus \mathcal{K}_{n-1}$ , as

QC-QC

$$= \left\{ \begin{array}{l} W \in \mathcal{L}(\mathcal{H}_N^{A^{IO}F}), W \geq 0 \\ \text{s.t.} \left\{ \begin{array}{l} \text{Tr}_F W = \sum_{k_N \in \mathcal{N}} W_{(\mathcal{N} \setminus \{k_N\}, k_N)} \otimes \mathbb{1}^{A_{k_N}^O}, \\ \forall n = 1, \dots, N-1, \forall \emptyset \subsetneq \mathcal{K}_n \subsetneq \mathcal{N}, \\ \sum_{k_{n+1} \in \mathcal{N} \setminus \mathcal{K}_n} \text{Tr}_{A_{k_{n+1}}^I} W_{(\mathcal{K}_n, k_{n+1})} \\ = \sum_{k_n \in \mathcal{K}_n} W_{(\mathcal{K}_n \setminus \{k_n\}, k_n)} \otimes \mathbb{1}^{A_{k_n}^O}, \\ \sum_{k_1 \in \mathcal{N}} \text{Tr} W_{(\emptyset, k_1)} = 1, \\ \forall n = 1, \dots, N, \forall \mathcal{K}_{n-1} \subsetneq \mathcal{N}, \\ \forall k_n \in \mathcal{N} \setminus \mathcal{K}_{n-1}, W_{(\mathcal{K}_{n-1}, k_n)} \geq 0 \end{array} \right. \end{array} \right\}, \quad (24)$$

where the notation  $\mathcal{K}_n$  implicitly assumes that  $|\mathcal{K}_n| = n$ .

### C. Inclusion relations between the classes

Before considering the power of these different classes in the metrological task, let us first discuss the relation between these classes themselves.

First, note that in the rather trivial case of  $N = 1$ , all the classes introduced above coincide.

For  $N = 2$ , however, there are several non-trivial inclusion relations between the classes. Since parallel circuits can be seen as particular instances of circuits with fixed orders, QC-Par is a subclass of QC-FO, while QC-FO (for any order  $\pi$ ) is evidently a subclass of convQC-FO. It is easy to see that these inclusions are moreover strict. For  $N = 2$ , no dynamical control of the causal order can be considered since the choice of the first operation also forces the other operation to be second independently of what first operation is applied. Hence convQC-FO and QC-CC coincide, as do likewise QC-Sup and QC-QC. The quantum switch,  $W^{QS}$ , is in QC-QC yet not in convQC-FO which shows this inclusion is, however, strict. Finally, Gen includes, by construction, all other strategies and hence all these classes, and one can easily construct examples to show it strictly includes QC-QC [1, 34].

In the approach we use to compute the optimal QFI we optimize over  $W \in \text{Tr}_F \mathbf{W}$ . When the global future is traced out, the inclusion relations between the classes are slightly altered. As it turns out, in this case  $\text{Tr}_F \text{convQC-FO}$ ,  $\text{Tr}_F \text{QC-CC}$ ,  $\text{Tr}_F \text{QC-Sup}$  and  $\text{Tr}_F \text{QC-QC}$  all coincide since, in addition to there being no dynamical order for  $N = 2$ , the possibility of including circuits with causal superposition, which would distinguish QC-CC from QC-Sup and QC-QC, vanishes when  $F$  is traced out. The different inclusion relations for  $N = 2$ , both without and, resp., when tracing over  $F$  are

$$\text{QC-Par} \subsetneq \text{QC-FO} \subsetneq \text{convQC-FO} = \mathbf{QC-CC} \subsetneq \text{QC-Sup} = \mathbf{QC-QC} \subsetneq \text{Gen}, \quad (25)$$

$$\text{Tr}_F \text{QC-Par} \subsetneq \text{Tr}_F \text{QC-FO} \subsetneq \text{Tr}_F \text{convQC-FO} = \mathbf{Tr}_F \text{QC-CC} = \text{Tr}_F \text{QC-Sup} = \mathbf{Tr}_F \text{QC-QC} \subsetneq \text{Tr}_F \text{Gen}, \quad (26)$$

where the new classes we study here that have not previously been considered in quantum metrology are indicated in bold.

These are graphically represented in Figs. 4a and Fig. 4b, respectively.

For  $N \geq 3$ , the inclusion relations between QC-Par, QC-FO and convQC-FO remain the same, and Gen likewise includes all the other classes. However, due to the possibility of dynamical control of causal order that arises starting with  $N = 3$ , the equalities between convQC-FO and QC-CC, as well as between QC-Sup and QC-QC break down. QC-Sup becomes a strict subclass of QC-QC, as shown by the so-called Grenoble process [34] that gives an explicit example of a QC-QC that can readily be verified to not be in QC-Sup. Moreover, neither QC-CC nor QC-Sup is a subset of the other, since QC-CC contains circuits with dynamical control of causal order but without causal superposition (such as a ‘‘classical switch’’, where the first operation prepares a control system that determines

the order between the remaining ones [2, 34]), while QC-Sup contains circuits with causal superposition but without dynamical control of causal order (such as the generalization of the quantum switch to  $N$  operations).

In the case where  $F$  is traced out,  $\text{Tr}_F \text{convQC-FO}$  and  $\text{Tr}_F \text{QC-Sup}$  again coincide by construction, and are subclasses of  $\text{Tr}_F \text{QC-CC}$ , which allows for dynamical control of the causal order.  $\text{Tr}_F \text{QC-CC}$  is a strict subclass of  $\text{Tr}_F \text{QC-QC}$ , as again shown by the Grenoble process [34] which retains its properties even when  $F$  is traced out.

The different inclusion relations for  $N \geq 3$ , both without and, resp., when tracing over  $F$  are thus now

$$\text{QC-Par} \subsetneq \text{QC-FO} \subsetneq \text{convQC-FO} \subsetneq \text{QC-CC} \subsetneq \text{QC-QC} \subsetneq \text{Gen}, \quad (27)$$

$$\text{Tr}_F \text{QC-Par} \subsetneq \text{Tr}_F \text{QC-FO} \subsetneq \text{Tr}_F \text{convQC-FO} = \text{Tr}_F \text{QC-Sup} \subsetneq \text{Tr}_F \text{QC-CC} \subsetneq \text{Tr}_F \text{QC-QC} \subsetneq \text{Tr}_F \text{Gen}, \quad (28)$$

where QC-CC and QC-Sup are written on top of each other because they are incomparable. These relations are graphically represented in Figs. 5a and Fig. 5b, respectively.

We will now focus on different choices of external channel  $\mathcal{C}_\theta$ , in the cases  $N = 2, 3$ , in order to provide a quantitative comparison between the different classes previously introduced, and reassess the advantage of indefinite causal order for the quantum metrological task of channel parameter estimation.

## VI. RESULTS

A comparison of the classes QC-Par, QC-FO, QC-Sup and Gen through the metrological task presented above was already made in [19] for different choices of external channels  $\mathcal{C}_\theta$ , both for  $N = 2$  and  $N = 3$ . The classes QC-CC and QC-QC, however, are potentially more pertinent classes for understanding advantages arising from indefinite causal order. Indeed, the former corresponds to the most general strategies with a well-defined causal order, while the latter describes the most general strategies with a known physical implementation in terms of quantum-controlled circuits. Here, we study how these classes perform compared to those studied in [19] in the metrological task. First, we establish general hierarchies, based on the inclusion relations between the different classes presented above, which hold for any choice of channel  $\mathcal{C}_\theta$ . Then, we give a quantitative comparison and provide numerical hierarchies for particular choices of channel  $\mathcal{C}_\theta$ , using the SDP approach.

### A. General hierarchies

To start with, let us review the general hierarchies for  $N = 2, 3$  that follow from the inclusion relations given by Eqs. (25)–(28).

First, note that it is a general fact that (for any  $N$ ), by the convexity of the QFI,  $J_\theta^{\text{convQC-FO}}(C_\theta^{\otimes N}) = J_\theta^{\text{QC-FO}}(C_\theta^{\otimes N})$ . Hence, we will not consider  $J_\theta^{\text{convQC-FO}}(C_\theta^{\otimes N})$  explicitly any further. Secondly, as we discussed in Sec. IV B, not all QC-CCs are purifiable as QC-CCs; instead, the set of purifications of QC-CCs gives a subset of QC-QCs. We thus recall that, since the approach we use to calculate the QFI optimizes the QFI within the larger purifying class, it does not directly provide  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes N})$  but instead an upper bound on this quantity that we call  $J_{\theta, \text{purif}}^{\text{QC-CC}}(C_\theta^{\otimes N})$ .

#### 1. $N = 2$ case

Since for  $N = 2$ ,  $\text{convQC-FO} = \text{QC-CC}$  (see Eq. (25)), we have that  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 2}) = J_\theta^{\text{convQC-FO}}(C_\theta^{\otimes 2}) = J_\theta^{\text{QC-FO}}(C_\theta^{\otimes 2})$ . Furthermore, as written in Eq. (26), in the  $N = 2$  case  $\text{Tr}_F \text{QC-CC}$  coincides with  $\text{Tr}_F \text{QC-Sup}$  and  $\text{Tr}_F \text{QC-QC}$ , so it must be that  $J_{\theta, \text{purif}}^{\text{QC-CC}}(C_\theta^{\otimes 2}) = J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 2}) = J_\theta^{\text{QC-QC}}(C_\theta^{\otimes 2})$ . Combined with the other inclusion relations presented in Eq. (25) we thus obtain, for any quantum channel  $\mathcal{C}_\theta$ , the following hierarchy between the classes (where the QFI values for the new classes considered in this work are again highlighted in bold):

$$J_\theta^{\text{QC-Par}}(C_\theta^{\otimes 2}) \leq J_\theta^{\text{QC-FO}}(C_\theta^{\otimes 2}) = \mathbf{J_\theta^{\text{QC-CC}}}(C_\theta^{\otimes 2}) \leq \mathbf{J_{\theta, \text{purif}}^{\text{QC-CC}}}(C_\theta^{\otimes 2}) = J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 2}) = \mathbf{J_\theta^{\text{QC-QC}}}(C_\theta^{\otimes 2}) \leq J_\theta^{\text{Gen}}(C_\theta^{\otimes 2}). \quad (29)$$

As we will see in the examples below in Sec. VIB 1 (see

also [19]), the inequalities in this hierarchy are strict for at least

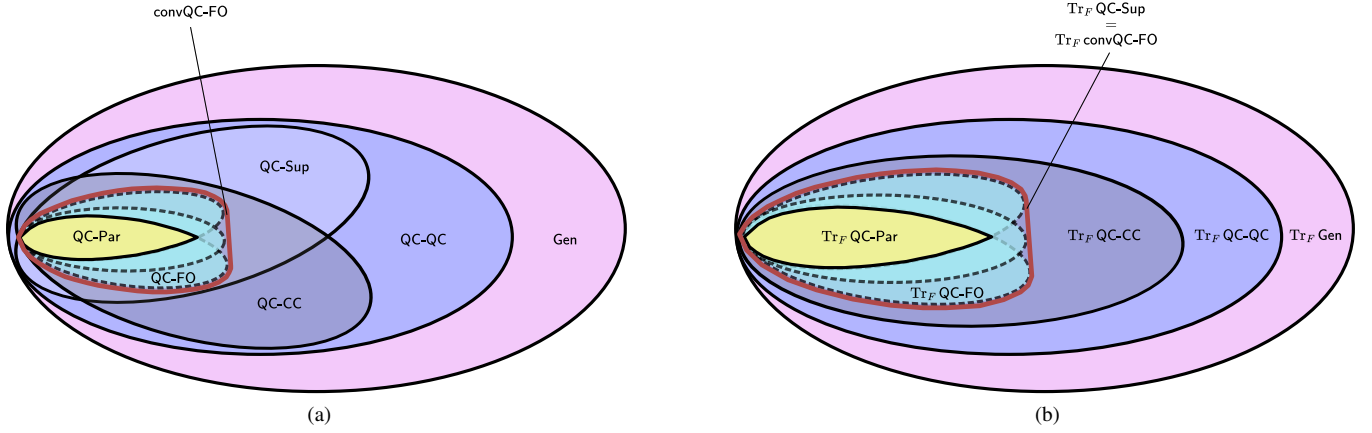


Figure 5. Relations between the different classes of strategies with and without definite causal order for  $N = 3$ , in the cases where the global future  $F$  is (a) not traced out; (b) traced out.

some choice of external channels  $\mathcal{C}_\theta$ , and so Eq. (29) cannot be strengthened further in general.

## 2. $N = 3$ case

For  $N = 3$ , as indicated in Eq. (28) we have  $\text{Tr}_F \text{QC-Sup} \subsetneq \text{Tr}_F \text{QC-CC}$ , which directly implies that  $J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 3}) \leq J_{\theta, \text{purif}}^{\text{QC-CC}}(C_\theta^{\otimes 3})$ . We likewise have trivially that  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 3}) \leq J_{\theta, \text{purif}}^{\text{QC-CC}}(C_\theta^{\otimes 3})$ . However, the relation between  $J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 3})$  and  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 3})$  is *a priori* unclear, and may in principle differ depending on  $\mathcal{C}_\theta$ . Combing this with the inclusion relations presented in Eq. (27), we obtain, for any quantum channel  $\mathcal{C}_\theta$ , the following hierarchies between the QFIs:

$$J_\theta^{\text{QC-Par}}(C_\theta^{\otimes 3}) \leq J_\theta^{\text{QC-FO}}(C_\theta^{\otimes 3}) \leq \underbrace{J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 3})}_{=?} \leq J_{\theta, \text{purif}}^{\text{QC-CC}}(C_\theta^{\otimes 3}) \leq J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 3}) \leq J_\theta^{\text{QC-QC}}(C_\theta^{\otimes 3}) \leq J_\theta^{\text{Gen}}(C_\theta^{\otimes 3}). \quad (30)$$

The question mark here indicates that, as discussed above, the value of  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 3})$  can only be bounded but not evaluated explicitly. The fact that  $J_\theta^{\text{QC-CC}}(C_\theta^{\otimes 3})$  and  $J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes 3})$  can in general not be compared is illustrated here by writing them on top of each other.

## B. Examples

We now look at some examples to illustrate the hierarchies previously introduced and to study the tightness of the inequalities

they contain for different families of channels. For each example considered, one can then conclude whether one can obtain an advantage using indefinite causal order strategies over definite causal order strategies.

### 1. Qubit depolarizing, Pauli, and thermalizing channels

Using the SDP approach for  $N = 1, 2, 3$  (qubit) depolarizing channels  $D_\theta$  as defined by Eq. (9), with Choi matrix  $D_\theta$ , we obtain the following equalities between the QFIs of the different classes, up to a numerical error of  $10^{-6}$ :

$$J_\theta^{\text{QC-Par}}(D_\theta^{\otimes N}) = J_\theta^{\text{QC-FO}}(D_\theta^{\otimes N}) = J_\theta^{\text{QC-CC}}(D_\theta^{\otimes N}) = J_\theta^{\text{QC-Sup}}(D_\theta^{\otimes N}) = J_\theta^{\text{QC-QC}}(D_\theta^{\otimes N}) = J_\theta^{\text{Gen}}(D_\theta^{\otimes N}) = \frac{3N}{(1-\theta)(1+3\theta)}. \quad (31)$$

The optimal strategy we found using the SDP approach

corresponds to the  $N$ -Choi circuit, which directly generalizes



the 2-Choi circuit we described in Sec. III B. Indeed, solving the SDP optimization problem for the largest class Gen we found precisely the strategy  $\tilde{W} = \frac{1}{2^N} \mathbb{1}_{A_N^{\otimes 2}}$ , which one can purify into the  $N$ -Choi circuit (see Appendix A 1). Since the  $N$ -Choi circuit is a parallel circuit, we conclude that there is no advantage of indefinite causal order strategies over definite causal order strategies for  $N = 1, 2, 3$  depolarizing channels  $\mathcal{D}_\theta$ , contrary to what was claimed in [12]. We conjecture that the above equalities hold for any  $N$  and any probe-dimension  $d$  (indeed, we verified this for  $N = 1, 2$  also in the qutrit case  $d = 3$ ), and that the corresponding optimal strategy remains the  $N$ -Choi circuit.

One can extend these results (in the qubit case) by considering more general Pauli channels  $\mathcal{P}_\vartheta$ , defined as

$$\mathcal{P}_\vartheta(\rho) = \vartheta\rho + (1 - \vartheta)(\alpha\sigma_x\rho\sigma_x + \beta\sigma_y\rho\sigma_y + \gamma\sigma_z\rho\sigma_z), \quad (32)$$

where  $\sigma_x, \sigma_y$  and  $\sigma_z$  are the Pauli matrices, and  $\alpha, \beta, \gamma \in [0, 1]$  such that  $\alpha + \beta + \gamma = 1$ . This type of channel describes many usual noise channels considered in quantum information theory, such as depolarizing, phase-flip and bit-flip channels. Here, we denoted the metrological parameter by  $\vartheta$ , rather than  $\theta$ , to give the Pauli channel expression a symmetric form. By considering the change of variable  $\vartheta \rightarrow \frac{1+3\theta}{4}$ , and the choice of parameters  $\alpha = \beta = \gamma = 1/3$ , one recovers the depolarizing channel as defined by Eq. (9).

The same equalities given in Eq. (31) between the QFIs of all the different classes were again obtained for Pauli channels with  $N = 1, 2, 3$ :  $\forall (\alpha, \beta, \gamma), J_\vartheta^{\mathbf{W}}(\mathcal{P}_\vartheta^{\otimes N}) = \frac{N}{\vartheta(1-\vartheta)}$ , for all classes  $\mathbf{W}$  considered in this work. As a sanity check one can verify that, with the change of variable  $\vartheta \rightarrow \frac{1+3\theta}{4}$ , we recover the optimal QFI of Eq. (31), and an optimal strategy is again the  $N$ -Choi circuit in that case (see again Appendix A 1). Note that this was already proven in Ref. [52] for the class QC-Par. We hence see that no advantage of indefinite causal order can be obtained in the metrological task for Pauli channels. As for depolarizing channels, we conjecture that the above equalities holds for any  $N$ , and that the corresponding optimal strategy remains the  $N$ -Choi circuit.

The exact same study for the thermalizing channels defined by Eq. (10) leads again to an identical conclusion. For

$$J_\theta^{\text{QC-Par}}(A_\theta^{\otimes 2}) < J_\theta^{\text{QC-FO}}(A_\theta^{\otimes 2}) = J_\theta^{\text{QC-CC}}(A_\theta^{\otimes 2}) < J_{\theta, \text{purif}}^{\text{QC-CC}}(A_\theta^{\otimes 2}) = J_\theta^{\text{QC-Sup}}(A_\theta^{\otimes 2}) = J_\theta^{\text{QC-QC}}(A_\theta^{\otimes 2}) = J_\theta^{\text{Gen}}(A_\theta^{\otimes 2}) \quad (35)$$

(except at  $u = 0, 1$ , when the hierarchy collapses completely). The numerical values of the QFI are given in Appendix C. In this case all the inequalities in Eq. (29) are found to be strict—i.e., we obtain strict separations between the corresponding classes—except for the last one—i.e., more general strategies in Gen do not outperform causal superpositions. Since QC-CC constitutes the set of causally separable processes for  $N = 2$ ,

$N = 1, 2, 3$  thermalizing channels  $\mathcal{T}_\theta$ , we numerically found the optimal QFI to be  $N \frac{\epsilon^2 e^{-\epsilon/\theta}}{\theta^4 (1+e^{-\epsilon/\theta})^2}$  for all the classes  $\mathbf{W}$  considered in this work (see Appendix A 2). We conclude that, contrary to the advantage claimed in [11], no advantage of indefinite causal order strategies can be found for  $N = 1, 2, 3$ . The optimal circuit was found to be the parallel circuit generalizing that presented in Fig. 3b for  $N = 2$ . We also conjecture that these results hold for any value of  $N$ , and that the optimal strategy still consists of  $N$  thermalizing channels used in parallel, with any input state.

Since no advantage could be obtained for depolarizing, Pauli, or thermalizing channels, we consider further examples of channels in the following subsections.

## 2. Rotation composed with an amplitude damping channel

Let us consider a  $z$ -rotation on the Bloch sphere of an angle  $\theta$  that we would like to estimate, and which is described by the unitary  $U_z(\theta) = e^{-i\theta\sigma_z/2}$ . As it turns out, this choice of channel does not display any gap between the QFI of the different classes. To see a gap, following Ref. [19] we then compose the rotation with an amplitude damping channel  $\mathcal{A}$ , defined by its two Kraus operators:

$$\begin{aligned} E_0 &= |0\rangle\langle 0| + \sqrt{1-u}|1\rangle\langle 1|, \\ E_1 &= \sqrt{u}|0\rangle\langle 1|, \end{aligned} \quad (33)$$

where  $u$  is the decay parameter. The composition of the rotation and the amplitude damping channel then defines the channel we consider,

$$\mathcal{A}_\theta = \mathcal{A} \circ \mathcal{U}_z(\theta), \quad (34)$$

with  $\mathcal{U}_z(\theta) : \rho \mapsto U_z(\theta)\rho U_z^\dagger(\theta)$ , and we will write its corresponding Choi matrix as  $A_\theta$ . It is easy to see that when optimizing over any of the classes under consideration here, the QFI will not depend on the specific value of  $\theta$ ,<sup>6</sup> so that one can fix  $\theta$  to any given value.

For  $N = 2$ , solving numerically the SDP to calculate the QFI for the different classes we find, for a range of values of  $u$ , the following hierarchy (where equalities are obtained up to an error of  $10^{-6}$ )

Eq. (35) provides a numerical proof that, for the quantum channels  $\mathcal{A}_\theta$ , causally nonseparable processes outperform causally separable ones.

For the  $N = 3$  scenario, by conducting similar numerical studies we found (again, except for  $u = 0, 1$ ) the following hierarchy for these channels:

<sup>6</sup> Indeed if one changes  $\theta \rightarrow \theta + \theta'$ , one can write  $\mathcal{A}_{\theta+\theta'} = \mathcal{A}_\theta \circ \mathcal{U}_z(\theta')$ , and “correct” the first rotation (with angle  $\theta'$ ) by incorporating its inverse in the process matrices that we optimize over.

$$J_{\theta}^{\text{QC-Par}}(A_{\theta}^{\otimes 3}) < J_{\theta}^{\text{QC-FO}}(A_{\theta}^{\otimes 3}) \leq \underbrace{J_{\theta}^{\text{QC-CC}}(A_{\theta}^{\otimes 3})}_{=?} \leq J_{\theta}^{\text{QC-Sup}}(A_{\theta}^{\otimes 3}) = J_{\theta, \text{purif}}^{\text{QC-CC}}(A_{\theta}^{\otimes 3}) < J_{\theta}^{\text{QC-QC}}(A_{\theta}^{\otimes 3}) < J_{\theta}^{\text{Gen}}(A_{\theta}^{\otimes 3}). \quad (36)$$

The numerical results are again provided in Appendix C. In this case all the inequalities in Eq. (30) are found to be strict except between  $J_{\theta}^{\text{QC-Sup}}(A_{\theta}^{\otimes 3})$  and  $J_{\theta, \text{purif}}^{\text{QC-CC}}(A_{\theta}^{\otimes 3})$ , which were found to coincide here.

From Eq. (36), one sees that QC-QC strategies strictly outperform both QC-CC and QC-Sup strategies. Although the optimal QC-QC strategy could be obtained from the primal SDP problem, its form is rather complicated so we omit it here. Since in the  $N = 3$  case the QC-CC class once again coincides with the set of causally separable processes [50], Eq. (36) gives a numerical proof of the strict advantage of indefinite causal order strategies over definite causal order strategies in the metrological task we study. The fact that QC-QCs strictly outperform QC-Sups shows that not only causal superposition, but also the dynamical control of the causal order, are resources for this metrological task. Eq. (36) also shows that QC-QC strategies are outperformed by more general causally indefinite strategies in that case. Yet, it remains unclear whether such strategies can be described as generalized quantum circuits, or are even consistent with the laws of physics.

As we already discussed earlier, the QFI for QC-CC could only be upper-bounded, but not evaluated explicitly. We found moreover that the optimal process matrix  $\tilde{W}_{\text{QC-CC}}^{(\text{opt})} \in \mathcal{L}(\mathcal{H}_{\mathcal{N}}^{A_{\theta}^{\otimes 3}})$ , for a given value of  $\theta$ , obtained numerically by solving the primal SDP problem belonged to  $\text{Tr}_F \text{QC-Sup} \subsetneq \text{Tr}_F \text{QC-CC}$ . The purification  $W_{\text{QC-CC}}^{(\text{opt})} \in \mathcal{L}(\mathcal{H}_{\mathcal{N}^{\otimes 3}}^{A_{\theta}^{\otimes 3}})$  of this strategy is thus contained in QC-Sup and, in principal, one may have  $W_{\text{QC-CC}}^{(\text{opt})} \notin \text{QC-CC}$  (cf. Fig. 5a). This is, in fact, the case for the process we found numerically in this case. This implies that using purifications to extend strategies that are optimal over  $\text{Tr}_F \text{QC-CC}$  does not provide much insight about the optimal QC-CC here. One may hope that alternative methods could be found to extend the optimized process over  $\text{Tr}_F \text{QC-CC}$  on  $F$ , but we could not find such a procedure giving a process in QC-CC as desired. It thus remains unclear whether the upper bound  $J_{\theta, \text{purif}}^{\text{QC-CC}}(A_{\theta}^{\otimes 3})$  is tight or not in this case.

### 3. Rotation composed with random channels

Finally, to get some insight on how typical the advantages of the various classes are, we also briefly comment on the hierarchies obtained using random channels defined in the same way as was done in Ref. [19]. These channels are defined in analogy with the channel  $\mathcal{A}_{\theta}$ , replacing the amplitude damping channel by a random noise channel sampled from a set of

CP and trace preserving maps.<sup>7</sup> We again check for equalities up to a numerical precision of  $10^{-6}$ .

For the case of  $N = 2$ , for 876 of the 1000 random channels sampled we obtained the hierarchy given by Eq. (29), with all inequalities holding strictly. Most of the remaining cases (120 out of 1000), corresponded to the hierarchy of Eq. (35), where general strategies from Gen provided no advantage over QC-QC and QC-Sup strategies. The other 4 cases differed from Eq. (29) or (35) solely in that QC-Sup provided no advantage over QC-FO.

For  $N = 3$ , we performed a similar test and found that, for 967 of the 1000 random channels we sampled, the hierarchy obtained was the same as the one already obtained for a rotation composed with an amplitude damping channel (i.e.,  $\mathcal{A}_{\theta}$  as defined in Eq. (34)) and given by Eq. (36). The 33 remaining cases corresponded to hierarchies different from Eq. (36) in the sense that either QC-QC provided no advantage over QC-Sup while Gen still strictly outperformed QC-QC, or both QC-QC and Gen performed no better than QC-Sup. Overall, almost all of the random channels tested provided a strict advantage of QC-QC strategies over QC-Sup strategies and hence over any strategy with a definite causal order.

## VII. CONCLUSIONS

Causal indefiniteness has attracted significant recent attention as a novel resource in quantum information, often based on applications of the canonical quantum switch [2]. A number of recent works have, in particular, attempted to use the quantum switch to provide advantages in quantum metrology [11–14, 16–18, 20–23] in scenarios where the goal is, broadly speaking, to maximize the precision of estimation of an unknown parameter  $\theta$  carried by a quantum channel  $\mathcal{C}_{\theta}$ , given  $N$  queries to this quantum channel. In this paper we critically reassessed some of these claims. Taking the results of [11, 12] for illustration, we argued that the advantages exhibited cannot be interpreted as being due to indefinite causal order, since the quantum switch was compared only to a single causally ordered protocol, rather than the entire class of causally definite strategies. As we show, one can in fact obtain better performance for the channels they consider using a simple parallel probing strategy than with the quantum switch. The quantum switch is thus, contrary to the initial claims, not a resource for metrology for the depolarizing and thermalizing channels considered in [11, 12]. While we examined explicitly these two results, our general criticism applies also to many of the other claimed advantages [14, 16–18, 23].

<sup>7</sup> We generated these random channels numerically using the function “RandomSuperoperator” from the MATLAB library QETLAB [53]. They were chosen to be qubit channels with 2 Kraus operators.

To provide a more rigorous study of metrological advantages arising from causal indefiniteness, we considered the framework recently introduced in [15, 19] that allows one to optimize the metrological performance over whole classes of strategies with different properties using SDPs. We extended this framework to include the most general known classes of physically realizable circuits with definite (QC-CC) and indefinite (QC-QC) causal order [34]. These classes notably allow for dynamical control of the order in which the channels are probed, or queried, allowing this to be determined on the fly, either incoherently (for QC-CCs) or coherently in a quantum superposition (for QC-QCs). We were thus able to study whether such dynamical control is useful in metrology, and to compare a wide range of physically pertinent classes of strategies: parallel (QC-Par), fixed order (QC-FO), and classically controlled (QC-CC) strategies, all which are causally definite, and causal superpositions (QC-Sup), quantum control (QC-QC) and general strategies (Gen), which also include causally indefinite processes.

We studied the performance of these strategies with two and three queries to the channels  $C_\theta$  for several families of channels: general Pauli channels, thermalizing channels, a rotation composed with an amplitude damping channel, and for a rotation composed with different random channels. For Pauli and thermalizing channels, we found that indefinite causal order provides no advantage, while for the other families QC-Sup, QC-QC and Gen all provided advantages over causally definite strategies. With three queries, in particular, we saw a strict hierarchy, with QC-QCs providing an advantage over QC-Sups which, in turn, outperformed QC-CCs. We thus extended the hierarchies obtained in [19] and showed that causal indefiniteness (and, indeed, dynamical quantum control of causal order) can provide advantages in metrology over the most general causally definite strategies.

As we saw, causally indefinite strategies are useful when dealing with some quantum channels, but for other families they provided no advantage. It would be interesting to understand further when advantages can be obtained, and what properties of the channels  $C_\theta$  are behind such differences. All the channels we studied here were non-unitary, which may appear a key omission given the prevalence of unitary channels in quantum metrological problems. The recent results of [54] show, however, that QC-QCs cannot provide advantages over QC-FOs in any task (even beyond metrology) in which one queries a single unitary  $N$  times. In the type of metrological task we considered here they hence cannot provide an advantage. Nonetheless, in more complex settings where several unitary channels need to be probed (as considered for instance in [13]), advantages may still be possible. It would nonetheless be interesting to also see whether the class Gen of general strategies, or some pertinent subclass (e.g., that of purifiable processes [42]) can lead to advantages for unitary channels using causally indefinite strategies, even if the physical interpretation of such strategies remains unclear. More generally, it would be interesting to consider the utility of causally indefinite strategies beyond the simple metrological task we consider here, for example looking at multi-parameter estimation, where more than one parameter must be estimated, or multi-channel

parameter estimation, where different channels that depend on the same parameter must be probed.

Further work also remains to improve the approach used to compute the optimal QFI for QC-CC, which is not closed under purifications, either by finding an alternative approach that computes exactly the optimal QFI over the class, or by finding better upper bounds on it. Even if the approach here was sufficient for our purposes and indeed allowed us to show some advantages of QC-QC over QC-CC and QC-Sup in particular, this would help to clarify the hierarchy between QC-CC and QC-Sup when  $N \geq 3$ .

Finally, one may wonder about the physical relevance of the class QC-Sup, both in and beyond quantum metrology. Causally indefinite strategies in QC-Sup exploit only “static” causal superpositions as in the quantum switch (as opposed to dynamical control of causal order). A recent work found that, in the asymptotic regime, causal superpositions provide no advantages over fixed order strategies [21]. It would hence be interesting to understand whether QC-QCs, which allow for dynamical control of the causal order, could provide asymptotic advantages in such a setting. More generally, it seems that QC-Sup fails to fully capture all processes with “static” causal superpositions. Indeed, already by modifying slightly the description of the quantum switch and considering that the final control system is given as an input to a third operation, rather than to the future space  $\mathcal{H}^F$ , gives a process that formally is not in QC-Sup yet conceptually is identical to the quantum switch. This problem of formalizing more rigorously the distinction between static and dynamical causal order is the subject of a future work in preparation [55].

## ACKNOWLEDGMENTS

We thank Qiushi Liu and Yuxiang Yang for helpful discussions and comments on this work.

This research was funded in part by l’Agence Nationale de la Recherche (ANR) projects ANR-15-IDEX-02 and ANR-22-CE47-0012, and the PEPR integrated project EPiQ ANR-22-PETQ-0007 as part of Plan France 2030. For the purpose of open access, the authors have applied a CC-BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

## Appendix A: Processes with definite causal order outperforming the quantum switch

### 1. Depolarizing channels

Ref. [12] considered depolarizing channels  $D_\theta$  of the form of Eq. (9) and compared the performance of the quantum switch, versus a simple use of the channels in sequence, in estimating the parameter  $\theta$ .

For  $N = 2$  uses of the channels, the quantum switch is described by the process matrix  $W^{\text{QS}}$  of Eq. (8), while the process that prepares some (pure) input state

$|\psi_{\text{in}}\rangle$  and directly applies the two copies of  $\mathcal{D}_\theta$  one after the other is described as the process matrix  $W^{\text{Seq}} = |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^{A_1^I} \otimes |\mathbb{1}\rangle\langle\mathbb{1}|^{A_1^O A_2^I} \otimes |\mathbb{1}\rangle\langle\mathbb{1}|^{A_2^O F}$ . From these process matrices one can explicitly write the output state of the circuit in each case, obtained as  $D_\theta^{\otimes 2} * W^{\text{QS}}$  and  $D_\theta^{\otimes 2} * W^{\text{Seq}}$ , respectively. As these have relatively simple expressions,<sup>8</sup> for which an eigendecomposition can easily be found, one can calculate the QFI directly using Eq. (2). For the case of qubit ( $d = 2$ ) depolarizing channels, we thus find<sup>9</sup>

$$J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}}) = \frac{8(1 + 3\theta^2)}{(1 - \theta)(1 + 3\theta)(3 + 2\theta + 3\theta^2)}, \quad (\text{A1})$$

$$J_\theta(D_\theta^{\otimes 2} * W^{\text{Seq}}) = \frac{4\theta^2}{1 - \theta^4}, \quad (\text{A2})$$

for any input state  $|\psi_{\text{in}}\rangle$  in both circuits, in agreement with [12]. One can then show that  $J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}}) > J_\theta(D_\theta^{\otimes 2} * W^{\text{Seq}})$ ,  $\forall \theta \in [0, 1)$ , from which Ref. [12] concluded that the quantum switch provides a metrological advantage with respect to the sequential strategy.

Consider now the ‘‘2-Choi circuit’’ of Fig. 3a.<sup>10</sup> Since it consists in two independent uses of the Choi circuit, whose output state is simply (up to a proportionality factor of  $\frac{1}{d}$ ) the Choi matrix of the probed channel, by the additivity of the QFI one has  $J_\theta(D_\theta^{\otimes 2} * W^{2\text{-Choi}}) = 2J_\theta(\frac{1}{d}D_\theta) = 2J_\theta(\theta|\phi^+\rangle\langle\phi^+| + (1 - \theta)\frac{\mathbb{1}}{d^2})$ . Using Eq. (2), for the case of qubits ( $d = 2$ ) we find

$$J_\theta(D_\theta^{\otimes 2} * W^{2\text{-Choi}}) = \frac{6}{(1 - \theta)(1 + 3\theta)}. \quad (\text{A3})$$

It is then straightforward to show that  $J_\theta(D_\theta^{\otimes 2} * W^{2\text{-Choi}}) > J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}})$ ,  $\forall \theta \in [0, 1)$ , as can also be seen in Fig. 6. Hence, one cannot claim that the quantum switch has any advantage over general definite causal order strategies in this case.

<sup>8</sup> Specifically, one finds  $D_\theta^{\otimes 2} * W^{\text{QS}} = \theta^2 |+\rangle\langle+|^c \otimes |\psi\rangle\langle\psi|^t + 2\theta(1 - \theta) |+\rangle\langle+|^c \otimes \frac{\mathbb{1}^t}{d} + (1 - \theta)^2 \frac{\mathbb{1}^c}{2} \otimes \frac{\mathbb{1}^t}{d} + (1 - \theta)^2 \frac{|0\rangle\langle 1|^c + |1\rangle\langle 0|^c}{2d^2} \otimes |\psi\rangle\langle\psi|^t$  and  $D_\theta^{\otimes 2} * W^{\text{Seq}} = \theta^2 |\psi\rangle\langle\psi| + (1 - \theta^2) \frac{\mathbb{1}}{d}$  (the latter could of course have been obtained more directly).

<sup>9</sup> One could similarly obtain the (quite cumbersome) expression of the QFI for any dimension  $d$ , and recover the result given in [12]. The facts that the QFI does not depend on the (pure) input states and that  $J_\theta(D_\theta^{\otimes 2} * W^{\text{QS}}) > J_\theta(D_\theta^{\otimes 2} * W^{\text{Seq}})$  hold for any dimension.

<sup>10</sup> For completeness, the process matrix of the  $N$ -Choi circuit can be written as  $W^{N\text{-Choi}} = \bigotimes_{k=1}^N |\phi^+\rangle\langle\phi^+|^{A_k^I F_k} \otimes |\mathbb{1}\rangle\langle\mathbb{1}|^{A_k^O F'_k}$  (with a global future space  $\mathcal{H}^F = \bigotimes_{k=1}^N \mathcal{H}^{F_k F'_k}$ ).

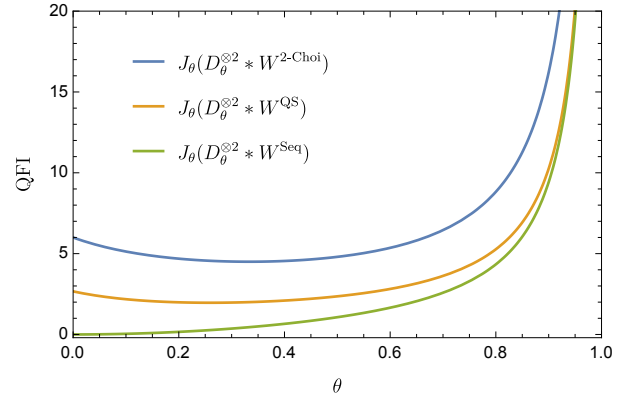


Figure 6. QFI of the output state of the quantum switch, of the sequential and of the 2-Choi strategies in the case of  $N = 2$  qubit ( $d = 2$ ) depolarizing channels, as defined in Eq. (9).

Note that for  $N$  qubit depolarizing channels, using again the additivity of the QFI, we readily obtain<sup>11</sup>

$$J_\theta(D_\theta^{\otimes N} * W^{N\text{-Choi}}) = \frac{3N}{(1 - \theta)(1 + 3\theta)}, \quad (\text{A4})$$

which is precisely the optimal QFI found numerically in Eq. (31) for  $N = 1, 2, 3$  qubit depolarizing channels.

## 2. Thermalizing channels

Rather than depolarizing channels, Ref. [11] considered the case with  $N = 2$  copies of a qubit thermalizing channel  $\mathcal{T}_\theta$ , as defined from the Kraus operators of Eq. (10) in the limit of infinite interaction time—in which case the channel reduces to the constant channel  $\rho \mapsto \tau_\theta := p_\theta|0\rangle\langle 0| + (1 - p_\theta)|1\rangle\langle 1|$ , with  $p_\theta = 1/(1 + e^{-\epsilon/\theta})$ .

When plugging the two copies of  $\mathcal{T}_\theta$  into quantum switch, the output state can be computed from the process matrix  $W^{\text{QS}}$  given in Eq. (8).<sup>12</sup> It is found to depend on the initial state  $|\psi_{\text{in}}\rangle$  of the target system; as in [11], we shall take it to be  $|\psi_{\text{in}}\rangle = |0\rangle$ . As for the output of the basic sequential composition of the two copies, it is trivially equal to the constant output state  $\tau_\theta$  of the channel itself. From an eigendecomposition of these two output states, using Eq. (2) with  $\partial_\theta \rho_\theta = (\partial_{p_\theta} \rho_\theta)(\partial_\theta p_\theta)$ ,

<sup>11</sup> More generally, for some arbitrary dimension  $d$  one has  $J_\theta(D_\theta^{\otimes N} * W^{N\text{-Choi}}) = \frac{(d^2 - 1)N}{(1 - \theta)(1 + (d^2 - 1)\theta)}$ . Note that Ref. [12] also considered a ‘‘quantum  $N$ -switch’’ that superposes two reverse orders for  $N$  operations. The value of the QFI obtained for that  $N$ -switch is again smaller than that of the  $N$ -Choi circuit given just above. (The analytical comparison being quite tedious, we verified this numerically for all values of  $N, d \leq 100$ ; the observed behavior convinces us that this remains true for all values of  $N, d$ .)

<sup>12</sup> Specifically, one finds  $T_\theta^{\otimes 2} * W^{\text{QS}} = \frac{\mathbb{1}^c}{2} \otimes \tau_\theta + \frac{|0\rangle\langle 1|^c + |1\rangle\langle 0|^c}{2} \otimes \tau_\theta |\psi\rangle\langle\psi| \tau_\theta$ .



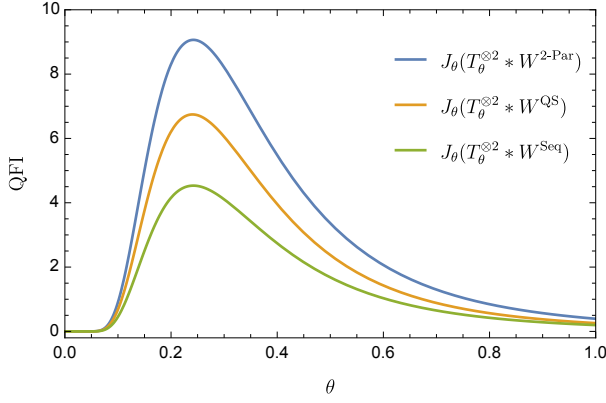


Figure 7. QFI of the output state of the quantum switch, of the sequential and of the parallel strategies in the case of  $N = 2$  thermalizing channels, as defined through Eq. (10), in the limit  $t \rightarrow \infty$ .

we find the corresponding QFIs to be

$$J_\theta(T_\theta^{\otimes 2} * W^{QS}) = \left( \frac{1}{p_\theta(1-p_\theta^2)} + \frac{1}{1-p_\theta} \right) (\partial_\theta p_\theta)^2, \quad (\text{A5})$$

$$J_\theta(T_\theta^{\otimes 2} * W^{Seq}) = J_\theta(\tau_\theta) = \left( \frac{1}{p_\theta} + \frac{1}{1-p_\theta} \right) (\partial_\theta p_\theta)^2, \quad (\text{A6})$$

which agrees with the result obtained in [11], and from which one clearly sees that  $J_\theta(T_\theta^{\otimes 2} * W^{QS}) > J_\theta(T_\theta^{\otimes 2} * W^{Seq})$  for all  $\theta > 0$ , thus suggesting some advantage for the quantum switch compared to the sequential strategy.

However, since one considers here two copies of the constant channel  $\mathcal{T}_\theta$ , it is more efficient to use them in parallel than in series, as in the process depicted in Fig. 3b (with its process matrix denoted  $W^{2\text{-Par}}$ ), so as to have access to two copies of the output state,  $\tau_\theta^{\otimes 2}$  rather than just  $\tau_\theta$ . By the additivity of the QFI, one gets

$$\begin{aligned} J_\theta(T_\theta^{\otimes 2} * W^{2\text{-Par}}) &= J_\theta(\tau_\theta^{\otimes 2}) = 2J_\theta(\tau_\theta) \\ &= 2 \left( \frac{1}{p_\theta} + \frac{1}{1-p_\theta} \right) (\partial_\theta p_\theta)^2. \end{aligned} \quad (\text{A7})$$

One can easily verify that  $J_\theta(T_\theta^{\otimes 2} * W^{2\text{-Par}}) > J_\theta(T_\theta^{\otimes 2} * W^{QS})$ , as also seen on Fig. 7. Thus, the parallel strategy outperforms the quantum switch for the present task, and no advantage of the quantum switch versus the general class of processes with definite causal order can be claimed here.

For  $N$  thermalizing channels, the parallel strategy (with process matrix  $W^{N\text{-Par}}$ ) gives

$$\begin{aligned} J_\theta(T_\theta^{\otimes N} * W^{N\text{-Par}}) &= N \left( \frac{1}{p_\theta} + \frac{1}{1-p_\theta} \right) (\partial_\theta p_\theta)^2 \\ &= N \frac{\epsilon^2 e^{-\epsilon/\theta}}{\theta^4 (1 + e^{-\epsilon/\theta})^2}. \end{aligned} \quad (\text{A8})$$

This expression is precisely the optimal QFI for all the classes considered in this work, found numerically for  $N = 1, 2, 3$

thermalizing channels; see Sec. VIB 1. (To obtain the numerical results we fixed  $\epsilon = 1$ ; the result for generic values of  $\epsilon$  is obtained by rescaling  $\theta \rightarrow \theta/\epsilon$ .)

## Appendix B: Computing the QFI and reconstructing the optimal process through SDP optimization

In this appendix we clarify further and complete the arguments for the approach we use—which was introduced in Refs. [15, 19], and which we only sketched in Sec. IV B—to compute the QFI and reconstruct the optimal process in the metrological task at hand.

### 1. Step-by-step description of the approach

Recall that our aim is to optimize the QFI of the state that is output by the composition of the  $N$   $\theta$ -dependent channels  $C_\theta$  (represented by their Choi matrix  $C_\theta$ ) with a process  $W$  from a given (compact) class  $\mathbf{W}$ , as given in Eq. (11) of the main text. That is,

$$J_\theta^{\mathbf{W}}(C_\theta^{\otimes N}) = \max_{W \in \mathbf{W}} J_\theta(C_\theta^{\otimes N} * W), \quad (\text{B1})$$

where the QFI  $J_\theta$  of the output state  $\rho_\theta = C_\theta^{\otimes N} * W$  can be computed, using Eq. (3), as

$$J_\theta(\rho_\theta) = 4 \min_{\{|\psi_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \sum_{j=1}^q |\dot{\psi}_{\theta,j}\rangle \langle \dot{\psi}_{\theta,j}| \right), \quad (\text{B2})$$

with the minimum being taken over all ensemble decompositions  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  of  $\rho_\theta$  (that are assumed to be continuously differentiable, with  $|\dot{\psi}_{\theta,j}\rangle = \partial_\theta |\psi_{\theta,j}\rangle$ ), of any given size  $q \geq \max_\theta \text{rank } \rho_\theta$ .

#### a. Restricting to pure processes

The first step in the approach is to restrict oneself to “pure” processes<sup>13</sup> of the form  $W = |w\rangle\langle w|$ .

For many classes  $\mathbf{W}$  of interest, restricting to pure processes does not induce any loss of generality. This is the case when (apart from the requirement that  $W \geq 0$ ) the conditions for a given process matrix  $W$  to belong to a class  $\mathbf{W}$  only impose constraints on its partial trace  $\text{Tr}_F W$  directly, rather than constraints on the full  $W$ . This is indeed the case for all the classes considered previously in [15, 19], as well as for the QC-QC class that we also consider in this work; see the characterizations of Eqs. (17), (18), (20), (21) and (24). To see that one can indeed safely restrict oneself to pure processes for such classes, note that any “mixed” process matrix

<sup>13</sup> Here by “pure” processes we simply refer to rank-1 processes, rather than to the more restrictive notion proposed in Ref. [42].

$W \in \mathcal{L}(\mathcal{H}^{A_N^{IO} F})$  in such a class  $\mathbf{W}$  can be “purified” into  $W' = |w'\rangle\langle w'| \in \mathcal{L}(\mathcal{H}^{A_N^{IO} F F'})$  by introducing some auxiliary system  $F'$ , such that  $\text{Tr}_{F'} W' = W$ . Then considering the global future system of  $W'$  to be  $FF'$ ,  $\text{Tr}_{FF'} W' = \text{Tr}_F W$  also satisfies the required constraints for  $W'$  to be in  $\mathbf{W}$  (recall that our definitions of the classes  $\mathbf{W}$  leaves open the specification of the future space). Noting, further, that the QFI of the output state can only decrease by partially tracing out  $F'$ , i.e.,  $J_\theta(C_\theta^{\otimes N} * |w'\rangle\langle w'|) \geq J_\theta(C_\theta^{\otimes N} * W)$ , it follows that it is indeed sufficient to consider pure processes to obtain the maximum in Eq. (B1).

For some other classes however (such as QC-CC class that we also consider here), the purification argument does not work so straightforwardly: purifying a process matrix may in general take us outside of the class being considered into a larger one. We still consider such purifications here, but the approach we use then only allows us in this case to obtain the optimal QFI in this larger class—which thus only provides an *upper bound* (that we denote by  $J_{\theta, \text{purif}}^{\mathbf{W}}(C_\theta^{\otimes N})$ ) on the QFI of the class we initially consider.

*b. Trading the optimization over decompositions  $\{|\psi_{\theta,j}\rangle\}_j$  of  $\rho_\theta$  for an optimization over decompositions  $\{|C_{\theta,j}\rangle\}_j$  of  $C_\theta^{\otimes N}$*

Once we restrict ourselves to pure processes  $W = |w\rangle\langle w|$ , it can be seen that for any fixed  $q \geq \max_\theta \text{rank } C_\theta^{\otimes N}$ , the ensemble decompositions of size  $q$  of  $C_\theta^{\otimes N}$  and the ensemble decompositions of size  $q$  of  $\rho_\theta$ , with  $\rho_\theta = C_\theta^{\otimes N} * |w\rangle\langle w|$ , are in one-to-one correspondence through the (“pure”<sup>14</sup>) link product with  $|w\rangle$ —in the sense that for any ensemble decomposition  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  of  $C_\theta^{\otimes N}$  of size  $q$ , there exists an ensemble decomposition  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  of  $\rho_\theta$  of the same size  $q$ , such that  $|\psi_{\theta,j}\rangle = |C_{\theta,j}\rangle * |w\rangle \forall j$ , and vice versa.

The forward statement is trivial (using  $\rho_\theta = C_\theta^{\otimes N} * W = \sum_{j=1}^q |C_{\theta,j}\rangle\langle C_{\theta,j}| * |w\rangle\langle w|$ ), and is obtained precisely by taking  $|\psi_{\theta,j}\rangle = |C_{\theta,j}\rangle * |w\rangle$ .

The reverse direction, on the other hand, is not so straightforward, because the link product is not trivial to invert [34]. To see that it also holds, note first that since we consider  $q \geq \max_\theta \text{rank } C_\theta^{\otimes N}$ , then there exists an ensemble decomposition  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  of  $C_\theta^{\otimes N}$ , of size  $q$ , which also gives (as above) an ensemble decomposition  $\{|\psi_{\theta,j}^0\rangle\}_{j=1}^q$  of  $\rho_\theta$  with  $|\psi_{\theta,j}^0\rangle = |C_{\theta,j}^0\rangle * |w\rangle$ . Consider then any other ensemble decomposition  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  of  $\rho_\theta$ , of size  $q$ . As  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  and  $\{|\psi_{\theta,j}^0\rangle\}_{j=1}^q$  are ensemble decompositions

of the same state and of the same size  $q$ , it follows from the HJW theorem [56] (see also the next subsection for a similar argument) that the  $|\psi_{\theta,j}\rangle$ ’s can be obtained from the  $|\psi_{\theta,j}^0\rangle$ ’s as  $|\psi_{\theta,j}\rangle = \sum_{k=1}^q U_{\theta,k,j} |\psi_{\theta,k}^0\rangle$ , where the  $U_{\theta,k,j}$ ’s are the coefficients of a  $q \times q$  unitary. We then have  $|\psi_{\theta,j}\rangle = \sum_{k=1}^q U_{\theta,k,j} |C_{\theta,k}^0\rangle * |w\rangle = |C_{\theta,j}\rangle * |w\rangle$  with  $|C_{\theta,j}\rangle = \sum_{k=1}^q U_{\theta,k,j} |C_{\theta,k}^0\rangle$ . It is then easy to verify that  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  thus defined indeed provides an ensemble decomposition of size  $q$  of  $C_\theta^{\otimes N}$ .

There is a delicate issue here relating to the differentiability of the decompositions [38]. We assume in this work that the channels  $C_\theta$  have a smooth enough dependency on  $\theta$ , so that starting from an ensemble decomposition of  $\rho_\theta$  in which the  $|\psi_{\theta,j}\rangle$ ’s are continuously differentiable, the decomposition of  $C_\theta^{\otimes N}$  constructed above can also be taken such that the  $|C_{\theta,j}\rangle$ ’s are continuously differentiable (we colloquially just write “smooth”).<sup>15</sup> As the correspondence between the two decompositions is through  $|\psi_{\theta,j}\rangle = |C_{\theta,j}\rangle * |w\rangle$ , we can then write  $|\dot{\psi}_{\theta,j}\rangle = |\dot{C}_{\theta,j}\rangle * |w\rangle$ , and we can turn the optimization over  $\{|\psi_{\theta,j}\rangle\}_{j=1}^q$  in Eq. (B2) into an optimization over the decompositions  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  of the channels’ operator  $C_\theta^{\otimes N}$  directly, so as to obtain

$$\begin{aligned} J_\theta(\rho_\theta) &= 4 \min_{\{|C_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \left( \sum_{j=1}^q |\dot{C}_{\theta,j}\rangle\langle \dot{C}_{\theta,j}| \right) * W \right) \\ &= 4 \min_{\{|C_{\theta,j}\rangle\}_{j=1}^q} \text{Tr} \left( \left( \sum_{j=1}^q |\dot{C}_{\theta,j}\rangle\langle \dot{C}_{\theta,j}| \right)^T \tilde{W} \right), \end{aligned} \quad (\text{B4})$$

where in the second line we evaluated the link product and defined  $\tilde{W} := \text{Tr}_F W$ . Note at this point that the QFI of  $\rho_\theta$  only depends on  $\tilde{W}$ , after tracing out the future space of the process  $W$ : indeed to actually obtain  $\rho_\theta$  one should reintroduce that future space by purifying  $\tilde{W}$  (see Sec. B 1 a above), but all purifications are equivalent in the sense that they give the same QFI.

*c. Trading the optimization over  $\{|C_{\theta,j}\rangle\}_j$  for an optimization over a unitary  $V_\theta$*

The next step is to write any ensemble decomposition  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  in terms of a fixed, “reference” decomposition  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  and a unitary matrix  $V_\theta$ .

Consider any fixed (smooth) ensemble decomposition  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  of  $C_\theta^{\otimes N}$  (of size  $q \geq \max_\theta \text{rank } C_\theta^{\otimes N}$ ), and some other (smooth) decomposition  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  that we want to optimize. For convenience, let us denote by  $\mathbf{C}_\theta^0 = (|C_{\theta,1}^0\rangle, \dots, |C_{\theta,q}^0\rangle)$  and  $\mathbf{C}_\theta = (|C_{\theta,1}\rangle, \dots, |C_{\theta,q}\rangle)$  the

<sup>14</sup> The link product, which was introduced in the “mixed case” (for matrices) in Eq. (6) of the main text, can also be defined in the “pure case” as follows [34]: for two vectors  $|a\rangle \in \mathcal{H}^{XY}$  and  $|b\rangle \in \mathcal{H}^{YZ}$  that are connected through a common system  $Y$ ,

$$\begin{aligned} |a\rangle * |b\rangle &:= (\mathbb{1}^{XZ} \otimes \langle\langle \mathbb{1} |^{YY}) (|a\rangle^{XY} \otimes |b\rangle^{YZ}) \\ &= (|a\rangle^{TY} \otimes \mathbb{1}^Z) (\mathbb{1}^X \otimes |b\rangle) \in \mathcal{H}^{XZ}. \end{aligned} \quad (\text{B3})$$

<sup>15</sup> Admittedly, the exact requirements that this imposes on the channels  $C_\theta$  should still be clarified (as for instance in Ref. [38], where the families of quantum channels  $\{C_\theta\}_\theta$  are required to have continuously differentiable decompositions and to be “piecewise regular”). This is however rather involved, and goes beyond the scope of our work.

$d_{A^{\otimes N}} \times q$  matrices whose column vectors are the elements of the two decompositions  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  and  $\{|C_{\theta,j}\rangle\}_{j=1}^q$ , respectively. Since these are decompositions of the same operator and of the same size, it again follows from the HJW theorem<sup>16</sup> [56] that the matrices  $\mathbf{C}_\theta^0$  and  $\mathbf{C}_\theta$  are related through  $\mathbf{C}_\theta = \mathbf{C}_\theta^0 V_\theta$ , for some  $q \times q$  unitary  $V_\theta$ . We again assume that for two smooth decompositions  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  and  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  of  $C_\theta^{\otimes N}$ , the unitary  $V_\theta$  that relates them can be taken to be smooth.

Conversely, given the fixed  $\mathbf{C}_\theta^0$  and any (smooth)  $q \times q$  unitary  $V_\theta$ , it is easily verified that the column vectors of  $\mathbf{C}_\theta = \mathbf{C}_\theta^0 V_\theta$  provide an alternative (smooth) ensemble decomposition of  $C_\theta^{\otimes N}$ .

Hence, once a reference decomposition  $\{|C_{\theta,j}^0\rangle\}_{j=1}^q$  is fixed, the optimization over  $\{|C_{\theta,j}\rangle\}_{j=1}^q$  in Eq. (B4) can be replaced by an optimization over such ( $\theta$ -dependent, smooth)  $q \times q$  unitaries  $V_\theta$ . Noting that  $\sum_{j=1}^q |\dot{C}_{\theta,j}\rangle\langle \dot{C}_{\theta,j}| = \dot{\mathbf{C}}_\theta \dot{\mathbf{C}}_\theta^\dagger$ , with  $\dot{\mathbf{C}}_\theta = \dot{\mathbf{C}}_\theta^0 V_\theta + \mathbf{C}_\theta^0 \dot{V}_\theta = (\dot{\mathbf{C}}_\theta^0 + \mathbf{C}_\theta^0 \dot{V}_\theta V_\theta^\dagger) V_\theta$ , we obtain

$$J_\theta(\rho_\theta) = 4 \min_{V_\theta} \text{Tr} \left( \left( (\dot{\mathbf{C}}_\theta^0 + \mathbf{C}_\theta^0 \dot{V}_\theta V_\theta^\dagger) (\dot{\mathbf{C}}_\theta^0 + \mathbf{C}_\theta^0 \dot{V}_\theta V_\theta^\dagger)^\dagger \right)^T \tilde{W} \right). \quad (\text{B5})$$

*d. Trading the optimization over  $V_\theta$  for an optimization over a Hermitian matrix  $h$*

Looking more closely at Eq. (B5), one then sees that the objective function only depends on the  $q \times q$  Hermitian matrix  $h_\theta = i \dot{V}_\theta V_\theta^\dagger$ .<sup>17</sup>

Conversely, from any (in general,  $\theta$ -dependent and smooth)  $q \times q$  Hermitian matrix  $h_\theta$ , one can define a (smooth)  $q \times q$  unitary  $V_\theta$  that satisfies  $h_\theta = i \dot{V}_\theta V_\theta^\dagger$ , by taking  $V_\theta = e^{-iH_\theta}$ , where  $H_\theta$  is an antiderivative (with respect to  $\theta$ ) of  $h_\theta$ .

Hence, the optimization over  $V_\theta$  in Eq. (B5) is equivalent to an optimization over  $q \times q$  Hermitian matrices  $h_\theta$ . Since the dependency of  $h_\theta$  on  $\theta$  no longer enters the objective function—e.g., the derivative of  $h_\theta$  with respect to  $\theta$  does not appear—from now on (and as we did in the main text) we simply drop the  $\theta$  subscript in  $h_\theta$ . Introducing the so-called “performance operator” [15, 47]

$$\Omega_\theta(h) = 4 \left( (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h) (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h)^\dagger \right)^T, \quad (\text{B6})$$

we now obtain

$$J_\theta(\rho_\theta) = \min_{h \in \mathbb{H}_q} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right). \quad (\text{B7})$$

*e. Swapping  $\max_{\tilde{W}}$  and  $\min_h$*

Injecting now Eq. (B7) into Eq. (B1), the QFI  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$  is obtained as the result of a max-min optimization problem—

with the maximization being now over  $\tilde{W} = \text{Tr}_F W \in \text{Tr}_F \mathbf{W}$ . As the objective function  $f(h, \tilde{W}) = \text{Tr}(\Omega_\theta(h) \tilde{W})$  is convex in  $h$  and concave (even linear) in  $\tilde{W}$ , and as  $\mathbf{W}$  (and hence  $\text{Tr}_F \mathbf{W}$ ) is assumed to be compact, then by Fan’s minimax theorem [48] one can swap  $\max_{\tilde{W}}$  and  $\min_h$ , and write

$$\begin{aligned} J_\theta^{\mathbf{W}}(C_\theta^{\otimes N}) &= \max_{\tilde{W} \in \text{Tr}_F \mathbf{W}} \min_{h \in \mathbb{H}_q} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right) \\ &= \min_{h \in \mathbb{H}_q} \max_{\tilde{W} \in \text{Tr}_F \mathbf{W}} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right). \end{aligned} \quad (\text{B8})$$

*f. Trading the inner ( $\max_{\tilde{W}}$ ) SDP problem for its dual*

At this point, the inner ( $\max_{\tilde{W}}$ ) optimization problem takes the standard form of a SDP problem. One can then consider the dual of this (primal) problem, which involves a minimization problem that we write for now in the generic form  $\min_{S \text{ s.t. } \dots} S_\theta$ : the precise form for the constraint “ $S \text{ s.t. } \dots$ ” will be specified, for each of the classes  $\mathbf{W}$  under consideration, in Sec. B 2 below.

The structure of the problem considered here ensures that strong duality holds (see below), so that the optimal values of the primal and dual SDP problems match. We can then trade the former for the latter in Eq. (B8), and obtain

$$J_\theta^{\mathbf{W}}(C_\theta^{\otimes N}) = \min_{h \in \mathbb{H}_q} \min_{S \text{ s.t. } \dots} S_\theta. \quad (\text{B9})$$

*g. Solving the global SDP problem*

It remains now to solve Eq. (B9). The constraints we generically wrote for now as “ $S \text{ s.t. } \dots$ ” are linear in  $\Omega_\theta(h)$ , hence quadratic in  $h$  (see Eq. (B6)). One can however use a standard trick in convex optimization, which is to use Schur complements to linearize the constraints in  $h$  (see again Sec. B 2 below). This allows us to write Eq. (B9) as an SDP problem, and solve it using standard tools.

We thus finally obtain the optimal QFI  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$ —or its upper bound  $J_{\theta, \text{purif}}^{\mathbf{W}}(C_\theta^{\otimes N})$ , see Sec. B 1 a above—for the class  $\mathbf{W}$  under consideration.

*h. Reconstructing the optimal process*

One may not only be interested in estimating the optimal QFI  $J_\theta^{\mathbf{W}}(C_\theta^{\otimes N})$  (or its upper bound  $J_{\theta, \text{purif}}^{\mathbf{W}}(C_\theta^{\otimes N})$ ), but also in recovering the optimal process  $\tilde{W}^{(\text{opt})} \in \text{Tr}_F \mathbf{W}$  that reaches it.

Our approach above provides us with the optimal value  $h^{(\text{opt})}$  of  $h$  for the minimization problem of Eq. (B9), or equivalently, for the min-max problem of Eq. (B8). What remains is to find the optimal value  $\tilde{W}^{(\text{opt})}$  of  $\tilde{W}$ . It is however not enough to just fix  $h = h^{(\text{opt})}$  and minimize the objective function over  $\tilde{W}$ , as for the fixed value  $h^{(\text{opt})}$ , there might be different values of  $\tilde{W}$  giving the same value of the objective function—which

<sup>16</sup> See Footnote 2 in the Supplemental Material of Ref. [15] for a more detailed argument.

<sup>17</sup> The fact that  $h_\theta$  is Hermitian follows directly from  $\partial_\theta (V_\theta V_\theta^\dagger) = 0$ .

may thus all be solutions to the min-max problem of Eq. (B8), but not necessarily of the max-min problem, which is what we need in order to be assured that  $\tilde{W}$  corresponds to a process achieving this QFI via the link of Eq. (B7).

To make sure that one also gets a solution of the max-min problem, one must further impose that the derivative over  $h$ , after fixing  $\tilde{W}$ , vanishes. That is, one needs to further impose that

$$\partial_h \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right) \Big|_{h=h^{(\text{opt})}} = 0, \quad (\text{B10})$$

which can be translated<sup>48</sup> here into the requirement that

$$\mathbf{C}_\theta^{0\dagger} \tilde{W}^T (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h^{(\text{opt})}) \text{ is Hermitian.} \quad (\text{B11})$$

We discuss and illustrate further the necessity for imposing Eq. (B10) in Sec. B3 below.

Hence, the optimal  $\tilde{W}^{(\text{opt})}$  can be obtained as the solution to the following SDP problem:

$$\begin{aligned} \tilde{W}^{(\text{opt})} = \arg \max_{\tilde{W}} \quad & \text{Tr} \left( \Omega_\theta(h^{(\text{opt})}) \tilde{W} \right) \\ \text{s.t.} \quad & \left\{ \begin{array}{l} \tilde{W} \in \text{Tr}_F \mathbf{W}, \\ \mathbf{C}_\theta^{0\dagger} \tilde{W}^T (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h^{(\text{opt})}) \in \mathbb{H}_q. \end{array} \right. \end{aligned} \quad (\text{B12})$$

From this, one can then take any purification of  $\tilde{W}^{(\text{opt})}$  to define the pure process  $W^{(\text{opt})} = |w\rangle\langle w|$  we started from, and which generates the optimal output state  $\rho_\theta^{(\text{opt})} = \mathbf{C}_\theta^{\otimes N} * W^{(\text{opt})}$ . Recall that when considering  $\mathbf{W}$  to be QC-Par, QC-FO, QC-Sup, QC-QC or Gen, it is guaranteed that  $W^{(\text{opt})}$  is in the desired class; for QC-CC however,  $\tilde{W} \in \text{Tr}_F \text{QC-CC}$  does in general not imply that its purification  $W^{(\text{opt})}$  is in QC-CC, and we may indeed only get an upper bound on  $J_\theta^{\text{QC-CC}}(\mathbf{C}_\theta^{\otimes N})$  (see Sec. B1a).

## 2. Explicit SDP formulations for the classes under consideration

In this section we shall provide the explicit formulations for the different SDP problems one encounters in the approach

<sup>48</sup> Here  $\partial_h f(h) = 0$  means that the derivative of the function  $f(h)$  with respect to all scalar variables that are used to define  $h$  must be null. Denoting by  $h_{k,l}^r$  and  $h_{k,l}^i$  the real and imaginary parts of the matrix elements  $h_{k,l}$ , respectively, this amounts to imposing that  $\partial_{h_{k,l}^r} f(h) = \partial_{h_{k,l}^i} f(h) = 0$  for all  $k, l$ .

For a Hermitian matrix  $h$ , all these partial derivatives can be encoded into a single matrix—which we take to formalize our notation  $\partial_h f(h)$ —with diagonal elements  $[\partial_h f(h)]_{k,k} = \partial_{h_{k,k}^r} f(h)$  and non-diagonal elements  $[\partial_h f(h)]_{k,l} = \frac{1}{2} (\partial_{h_{k,l}^r} f(h) + i \partial_{h_{k,l}^i} f(h))$ . (Indeed, using  $h_{k,l}^r = h_{l,k}^r$  and  $h_{k,l}^i = -h_{l,k}^i$  one gets  $\partial_{h_{k,l}^r} f(h) = [\partial_h f(h)]_{k,l} + [\partial_h f(h)]_{l,k}$  and  $\partial_{h_{k,l}^i} f(h) = -i([\partial_h f(h)]_{k,l} - [\partial_h f(h)]_{l,k})$  for  $k \neq l$ .)

With this definition of  $\partial_h f(h)$ , one gets in particular the nice properties that  $\partial_h \text{Tr}(Ah) = A$ ,  $\partial_h \text{Tr}(Bh^2) = Bh + hB$ , and hence  $\partial_h \text{Tr}[(F_0 + F_1 h)(F_0 + F_1 h)^\dagger G] = F_1^\dagger G(F_0 + F_1 h) + (F_0 + F_1 h)^\dagger G F_1$  for any matrices  $A, B, F_0, F_1$  and  $G$  that are independent from  $h$ . Writing  $\text{Tr}(\Omega_\theta(h) \tilde{W})$  in the latter form (through Eq. (B6)), we indeed find that Eq. (B10) translates into the condition of Eq. (B11)—which is also equivalent to that introduced in Algorithm 1 of Ref. [19].

described above, for the different classes of processes that we consider in this work.

SDP problems are commonly encountered in the field of quantum information. These can be written in different ways; the standard form that we will use here consists in writing them as the following pair of “primal” and “dual” problems [57]:

$$\begin{aligned} \max_{\mathcal{W}} \quad & \text{Tr}[\mathcal{A} \mathcal{W}] \\ \text{s.t.} \quad & \Phi(\mathcal{W}) = \mathcal{B}, \mathcal{W} \geq 0 \end{aligned} \quad (\text{B13})$$

and

$$\begin{aligned} \min_{\mathcal{S}} \quad & \text{Tr}[\mathcal{S} \mathcal{B}] \\ \text{s.t.} \quad & \Phi^\dagger(\mathcal{S}) \geq \mathcal{A}, \end{aligned} \quad (\text{B14})$$

where  $\mathcal{A}$  and  $\mathcal{W}$  are Hermitian operators in some space  $\mathcal{L}(\mathcal{H}^W)$ ,  $\mathcal{B}$  and  $\mathcal{S}$  are Hermitian operators in some space  $\mathcal{L}(\mathcal{H}^S)$ ,  $\Phi$  is a Hermiticity-preserving linear map from  $\mathcal{L}(\mathcal{H}^W)$  to  $\mathcal{L}(\mathcal{H}^S)$ , and  $\Phi^\dagger$  is its adjoint map (such that  $\text{Tr}[\Phi^\dagger(\mathcal{S}) \mathcal{W}] = \text{Tr}[\mathcal{S} \Phi(\mathcal{W})]$  for all  $\mathcal{W} \in \mathcal{L}(\mathcal{H}^W)$ ,  $\mathcal{S} \in \mathcal{L}(\mathcal{H}^S)$ ).

Weak duality ensures that the maximal value of the primal problem (B13) is no larger than the minimal value of the dual problem (B14). In all situations we consider here, strong duality also holds, meaning that the two optimal values actually coincide. This indeed follows from verifying the so-called Slater’s condition: specifically, this is ensured here by the fact that there always exists  $\mathcal{W} > 0$  such that  $\Phi(\mathcal{W}) = \mathcal{B}$  and a Hermitian  $\mathcal{S}$  such that  $\Phi^\dagger(\mathcal{S}) \geq \mathcal{A}$  [57].

The optimization problems we shall encounter will not readily be of the form above, as they will typically involve some constraints that are quadratic in  $h$ . One can however resort to a standard trick to linearize these. Namely, for any  $s \times s$  Hermitian matrix  $\mathbf{S}$ , and for any  $s \times r$  matrix  $\mathbf{C}$ , let us define the Hermitian matrix  $\mathbf{A}_{\text{Schur}}(\mathbf{S}, \mathbf{C})$  as a block matrix

$$\mathbf{A}_{\text{Schur}}(\mathbf{S}, \mathbf{C}) := \begin{pmatrix} \mathbb{1} & \mathbf{C}^\dagger \\ \mathbf{C} & \mathbf{S} \end{pmatrix} \quad (\text{B15})$$

(with  $\mathbb{1}$  being here the  $r \times r$  identity matrix, so as to match the dimensions). It then follows from the Schur Complement Lemma (see e.g. Theorem 1.12 in Ref. [58]), that

$$\mathbf{A}_{\text{Schur}}(\mathbf{S}, \mathbf{C}) \geq 0 \quad \text{if and only if} \quad \mathbf{S} \geq \mathbf{C} \mathbf{C}^\dagger. \quad (\text{B16})$$

We will use this “Schur complement trick” to linearize constraints of the form  $\tilde{\mathbf{S}} \otimes \mathbb{1}^{A_{k_N}^I} \geq \text{Tr}_{A_{k_N}^O} \Omega_\theta(h)$ , for some  $k_N$ . To write the right hand side in the form  $\mathbf{C} \mathbf{C}^\dagger$ , we will take

$$\begin{aligned} \mathbf{C} = \mathbf{C}_{k_N}(h) & := 2 \left( \langle j |^{A_{k_N}^O} (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h)^* \right)_{j=1}^{d_{A_{k_N}^O}} \\ & = 2 \left( \langle 1 |^{A_{k_N}^O} (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h)^* \quad \dots \quad \langle d_{A_{k_N}^O} |^{A_{k_N}^O} (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h)^* \right) \end{aligned} \quad (\text{B17})$$

(where  $*$  denotes here the complex conjugate). That is,  $\mathbf{C}_{k_N}(h)$  a block matrix composed by the blocks  $(\mathbb{1}^{A_{k_N}^O} \otimes \langle j |^{A_{k_N}^O}) (\dot{\mathbf{C}}_\theta^0 - i \mathbf{C}_\theta^0 h)^*$ , where  $\{ |j\rangle^{A_{k_N}^O} \}_{j=1}^{d_{A_{k_N}^O}}$



denotes an orthonormal basis of  $\mathcal{H}^{A_{k_N}^O}$  (with the  $d_{A_{k_N}^O}$  blocks being put on a line, so as to turn the  $d_{A_{k_N}^O} \times q$  matrix  $(\dot{C}_\theta^0 - iC_\theta^0 h)^*$  into a  $\left(\frac{d_{A_{k_N}^O}}{d_{A_{k_N}^O}}\right) \times (d_{A_{k_N}^O} q)$  matrix). With this, we define

$$\mathbf{A}_{\text{Schur}, k_N}(\tilde{\mathbf{S}}, h) := \mathbf{A}_{\text{Schur}}(\tilde{\mathbf{S}} \otimes \mathbb{1}^{A_{k_N}^I}, \mathbf{C}_{k_N}(h)), \quad (\text{B18})$$

so that

$$\begin{aligned} \tilde{\mathbf{S}} \otimes \mathbb{1}^{A_{k_N}^I} &\geq \text{Tr}_{A_{k_N}^O} \Omega_\theta(h) \\ \text{if and only if } \mathbf{A}_{\text{Schur}, k_N}(\mathbf{S}, h) &\geq 0. \end{aligned} \quad (\text{B19})$$

Note that  $\mathbf{A}_{\text{Schur}, k_N}(\mathbf{S}, h)$  is linear in  $h$ , as desired.

Referring to these general formulations, we will now recall how to derive and write the SDP problems for the QC-FO class, which were already given in Refs. [15, 19]. We will then present the case of the QC-CC and QC-QC classes, which follow quite similar lines. For completeness and ease of reference, we will also give explicitly the SDP formulations for the QC-Par, QC-Sup and Gen classes (whose derivation can also be found in Ref. [19]).<sup>19</sup>

#### a. SDP problems for the QC-FO class

Using Eq. (18), the primal problem

$$\begin{aligned} &\max_{\tilde{W} \in \text{Tr}_F \text{ QC-FO}} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right) \\ &= \max \text{Tr} \left( \left( \text{Tr}_{A_N^O} \Omega_\theta(h) \right) W_{(N)} \right) \\ &\text{s.t.} \begin{cases} \text{Tr } W_{(1)} = 1, \\ \forall n = 1, \dots, N-1, \text{Tr}_{A_{n+1}^I} W_{(n+1)} = W_{(n)} \otimes \mathbb{1}^{A_n^O}, \\ \forall n = 1, \dots, N, W_{(n)} \geq 0. \end{cases} \end{aligned} \quad (\text{B20})$$

<sup>19</sup> As mentioned in the main text, the authors of Ref. [19] made a simplifying assumption on the structure of the classes  $\mathbf{W}$  under consideration (or rather, on  $\text{Tr}_F \mathbf{W}$ ; see their Eq. (9)), which led them to unify the derivation of the SDP problems for the different classes in a different way from ours. Our QC-CC and QC-QC classes do however not satisfy this simplifying assumption, so we could not readily use their derivation for our purposes. Note also that the forms of the dual SDPs given in Refs. [15, 19] are slightly different—but clearly equivalent—from the ones we give below: they can be obtained from ours by effectively “renormalizing” the dual variables in  $\mathcal{S}$  (typically, taking  $S_\theta = \lambda$  and replacing all other  $S_{(\cdot)}$ ’s by  $\lambda S'_{(\cdot)}$ , for some new dual variables  $\lambda$  and  $S'_{(\cdot)}$ ; or in the Gen case, replacing  $S$  by  $\lambda S'$  with  $\text{Tr}[S'] = d_{A_N^I}$ ). The way one uses the Schur complement trick to linearize the last constraint is then also slightly different (as one should then use Schur complements of matrices of the form  $\begin{pmatrix} \lambda \mathbb{1} & \mathbf{C}^\dagger \\ \mathbf{C} & \mathbf{S}' \end{pmatrix}$ ), rather than as in Eq. (B15)).

involved in Eq. (B8) can be written in the standard form of Eq. (B13), with<sup>20</sup>

$$\begin{aligned} \mathcal{W} &= \bigoplus_{n=1}^N W_{(n)}, \\ \mathcal{A} &= \left( \bigoplus_{n=1}^{N-1} \mathbb{0}^{A_{\{1, \dots, n-1\}}^I A_n^I} \right) \oplus \text{Tr}_{A_N^O} \Omega_\theta(h), \\ \Phi(\mathcal{W}) &= \bigoplus_{n=0}^{N-1} \Phi_{(n)}(\mathcal{W}) \\ \text{with } \begin{cases} \Phi_{(0)}(\mathcal{W}) = \text{Tr } W_{(1)}, \\ \forall n = 1, \dots, N-1, \\ \Phi_{(n)}(\mathcal{W}) = \text{Tr}_{A_{n+1}^I} W_{(n+1)} - W_{(n)} \otimes \mathbb{1}^{A_n^O}, \end{cases} \\ \mathcal{B} &= 1 \oplus \bigoplus_{n=1}^{N-1} \mathbb{0}^{A_{\{1, \dots, n\}}^O}. \end{aligned} \quad (\text{B21})$$

where  $\mathbb{0}$  denotes a null matrix (in the space indicated by the superscript).

#### Introducing the dual variable

$$\mathcal{S} = \bigoplus_{n=0}^{N-1} S_n \in \bigoplus_{n=0}^{N-1} \mathcal{L}(\mathcal{H}^{A_{\{1, \dots, n\}}^O}) \quad (\text{B22})$$

(with  $\mathcal{L}(\mathcal{H}^{A_\emptyset^O}) = \mathbb{R}$  for  $n = 0$ ), the adjoint of the linear map

<sup>20</sup> Other (ultimately equivalent) choices are possible. E.g., given the redundancies in the way we wrote the constraints for  $\tilde{W} \in \text{Tr}_F \text{ QC-FO}$  (namely, that all  $W_{(n)}$ ’s are uniquely determined by  $W_{(N)}$ , and that their positive semidefiniteness simply follows from that of  $W_{(N)}$ ), one could just take  $\mathcal{W} = W_{(N)}$ ,  $\mathcal{A} = \text{Tr}_{A_N^O} \Omega_\theta(h)$ , with the same  $\Phi$  and  $\mathcal{B}$  as in Eq. (B21), after replacing  $W_{(n)}$  in  $\Phi$  by  $\frac{1}{d_n^O d_{n+1}^O \dots d_{N-1}^O} \text{Tr}_{A_n^O A_{\{n+1, \dots, N-1\}}^I A_N^I} W_{(N)}$ . The choice of  $\mathcal{W}$  and  $\mathcal{A}$  in Eq. (B21), however, allows us to follow nicely parallel derivations for the QC-CC and QC-QC classes; see Sec. B 2 b below.

$\Phi$  is easily found to be<sup>21</sup>

$$\Phi^\dagger(\mathcal{S}) = \bigoplus_{n=1}^N (\Phi^\dagger)_{(n)}(\mathcal{S})$$

$$\text{with } \begin{cases} \forall n = 1, \dots, N-1, \\ (\Phi^\dagger)_{(n)}(\mathcal{S}) = S_{n-1} \otimes \mathbb{1}^{A_n^I} - \text{Tr}_{A_n^O} S_n, \\ (\Phi^\dagger)_{(N)}(\mathcal{S}) = S_{N-1} \otimes \mathbb{1}^{A_N^I}, \end{cases} \quad (\text{B23})$$

so that the dual problem, Eq. (B14), gets here the specific form

$$\min_{\mathcal{S}} S_0$$

$$\text{s.t. } \begin{cases} \forall n = 1, \dots, N-1, \\ S_{n-1} \otimes \mathbb{1}^{A_n^I} \geq \text{Tr}_{A_n^O} S_n, \\ S_{N-1} \otimes \mathbb{1}^{A_N^I} \geq \text{Tr}_{A_N^O} \Omega_\theta(h). \end{cases} \quad (\text{B24})$$

To see that Slater's condition for strong duality is satisfied (see above), note that  $\mathcal{W} = \bigoplus_{n=1}^N \frac{\mathbb{1}^{A_{\{1, \dots, n-1\}}^{IO}}}{d_{A_{\{1, \dots, n\}}^I}}$  for instance

satisfies  $\Phi(\mathcal{W}) = \mathcal{B}$  and  $\mathcal{W} > 0$ , and that one can always rather trivially find a Hermitian  $\mathcal{S}$  such that  $\Phi^\dagger(\mathcal{S}) \geq \mathcal{A}$ .

At this point one may note that it is possible to replace the first  $N-1$  inequalities in the constraints of Eq. (B24) by equalities. Indeed, if  $\mathcal{S}$  is an optimal solution of Eq. (B24), then  $\mathcal{S}' = \bigoplus_{n=0}^{N-1} S'_n$ , with

$$S'_n = S_0 \frac{\mathbb{1}^{A_{\{1, \dots, n\}}^{IO}}}{d_{A_{\{1, \dots, n\}}^O}} + \sum_{m=1}^n [1 - A_m^O] S_m \otimes \frac{\mathbb{1}^{A_{\{m+1, \dots, n\}}^{IO}}}{d_{A_{\{m+1, \dots, n\}}^O}}$$

$$= S_n + \sum_{m=1}^n \left( S_{m-1} \otimes \mathbb{1}^{A_m^I} - \text{Tr}_{A_m^O} S_m \right) \otimes \frac{\mathbb{1}^{A_m^O A_{\{m+1, \dots, n\}}^{IO}}}{d_{A_{\{m, \dots, n\}}^O}}, \quad (\text{B25})$$

is also an optimal solution (such that  $S'_0 = S_0$ ), which satisfies  $S'_{n-1} \otimes \mathbb{1}^{A_n^I} = \text{Tr}_{A_n^O} S'_n$  for  $n = 1, \dots, N-1$  (as is easily checked using the first form for  $S'_n$  given above), and  $S'_{N-1} \otimes$

$\mathbb{1}^{A_N^I} \geq S_{N-1} \otimes \mathbb{1}^{A_N^I} \geq \text{Tr}_{A_N^O} \Omega_\theta(h)$  (as is easily checked using the second expression for  $S'_{N-1}$ , using the constraints of Eq. (B24) that the  $S_n$ 's satisfy).

To obtain the optimal QFI for the QC-FO class, it remains to insert the above dual problem into Eq. (B9). Using the Schur complement trick introduced above, one can replace the quadratic (in  $h$ ) constraint  $S_{N-1} \otimes \mathbb{1}^{A_N^I} \geq \text{Tr}_{A_N^O} \Omega_\theta(h)$  by the linear constraint

$$\mathbf{A}_{\text{Schur}, N}(S_{N-1}, h) \geq 0, \quad (\text{B26})$$

with  $\mathbf{A}_{\text{Schur}, N}$  as defined in Eq. (B18). This thus turns the calculation of the QFI into an SDP problem:

$$J_\theta^{\text{QC-FO}}(C_\theta^{\otimes N})$$

$$= \min_{h, \mathcal{S}} S_0$$

$$\text{s.t. } \begin{cases} \forall n = 1, \dots, N-1, \\ S_{n-1} \otimes \mathbb{1}^{A_n^I} = \text{Tr}_{A_n^O} S_n, \\ \mathbf{A}_{\text{Schur}, N}(S_{N-1}, h) \geq 0. \end{cases} \quad (\text{B27})$$

As explained in Sec. B 1 h, the optimal process  $\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{QC-FO}$  is recovered by solving the primal problem of Eq. (B20) after fixing the optimal value of  $h = h^{(\text{opt})}$  obtained from Eq. (B27) and adding the constraint that  $\mathbf{C}_\theta^{0\dagger} \tilde{W}^T (\mathbf{C}_\theta^0 - i\mathbf{C}_\theta^0 h^{(\text{opt})}) \in \mathbb{H}_q$ , as in Eq. (B12)—and with  $\tilde{W} = W_{(N)} \otimes \mathbb{1}^{A_N^O}$  (cf. Eq. (18)).

The optimal process with fixed order  $W^{(\text{opt})} = |w\rangle\langle w| \in \text{QC-FO}$  is then obtained by purifying  $\tilde{W}^{(\text{opt})}$ .

#### b. SDP problems for the QC-CC and QC-QC classes

The derivation of the SDP problems for the QC-CC and QC-QC classes follows very similar lines to those for the QC-FO class, given in the previous subsection. We will not repeat everything; instead, we just present in the following table how Eqs. (B20) to (B27) are to be modified for the two new classes we considered in this paper.<sup>22</sup>

<sup>21</sup> From the structure of  $\mathcal{W}$ , one can write  $\text{Tr}[\Phi^\dagger(\mathcal{S})\mathcal{W}] = \sum_{n=1}^N \text{Tr}[(\Phi^\dagger)_{(n)}(\mathcal{S})W_{(n)}]$  and  $\Phi^\dagger(\mathcal{S}) = \bigoplus_{n=1}^N (\Phi^\dagger)_{(n)}(\mathcal{S})$ . Furthermore, the adjoint map satisfies  $\text{Tr}[\Phi^\dagger(\mathcal{S})\mathcal{W}] = \text{Tr}[\mathcal{S}\Phi(\mathcal{W})] = \sum_{n=0}^{N-1} \text{Tr}[S_n \Phi_{(n)}(\mathcal{W})] = S_0 \text{Tr} W_{(1)} + \sum_{n=1}^{N-1} \text{Tr}[S_n (\text{Tr}_{A_{n+1}^I} W_{(n+1)})] - \sum_{n=1}^{N-1} \text{Tr}[S_n (W_{(n)} \otimes \mathbb{1}^{A_n^O})] = S_0 \text{Tr} W_{(1)} + \sum_{n=2}^N \text{Tr}[(S_{n-1} \otimes \mathbb{1}^{A_n^I})W_{(n)}] - \sum_{n=1}^{N-1} \text{Tr}[(\text{Tr}_{A_n^O} S_n)W_{(n)}] =$

$\sum_{n=1}^{N-1} \text{Tr}[(S_{n-1} \otimes \mathbb{1}^{A_n^I} - \text{Tr}_{A_n^O} S_n)W_{(n)}] + \text{Tr}[(S_{N-1} \otimes \mathbb{1}^{A_N^I})W_{(N)}] = \sum_{n=1}^N \text{Tr}[(\Phi^\dagger)_{(n)}(\mathcal{S})W_{(n)}]$ ; as this must hold for all  $W_{(n)}$ 's, one can directly identify the terms inside the traces, for each term in the sums.

<sup>22</sup> For clarification: in the QC-CC case,  $(k_1, \dots, k_n)$  for  $n = 0$  is understood as the empty set,  $\emptyset$  (as e.g. in Eq. (B32):  $\mathcal{S} = \bigoplus_{n=0}^{N-1} \bigoplus_{(k_1, \dots, k_n)} S_{(k_1, \dots, k_n)} = S_\emptyset \oplus \bigoplus_{n=1}^{N-1} \bigoplus_{(k_1, \dots, k_n)} S_{(k_1, \dots, k_n)}$ ). In the QC-QC case,  $\mathcal{K}_n \setminus k_n$  is to be understood as  $\mathcal{K}_n \setminus \{k_n\}$ .

QC-CC	QC-QC
<p>Primal problem: (using Eq. (23))</p> $\begin{aligned} & \max_{\tilde{W} \in \text{Tr}_F \text{ QC-CC}} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right) \\ & = \max \text{Tr} \left( \sum_{(k_1, \dots, k_N)} \left( \text{Tr}_{A_{k_N}^O} \Omega_\theta(h) \right) W_{(k_1, \dots, k_N)} \right) \\ & \text{s.t.} \begin{cases} \sum_{k_1} \text{Tr} W_{(k_1)} = 1, \\ \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ \sum_{k_{n+1}} \text{Tr}_{A_{k_{n+1}}^I} W_{(k_1, \dots, k_n, k_{n+1})} \\ \quad = W_{(k_1, \dots, k_n)} \otimes \mathbb{1}_{A_{k_n}^O}, \\ \forall n = 1, \dots, N, \forall (k_1, \dots, k_n), W_{(k_1, \dots, k_n)} \geq 0. \end{cases} \end{aligned} \quad (\text{B28})$	<p>Primal problem: (using Eq. (24))</p> $\begin{aligned} & \max_{\tilde{W} \in \text{Tr}_F \text{ QC-QC}} \text{Tr} \left( \Omega_\theta(h) \tilde{W} \right) \\ & = \max \text{Tr} \left( \sum_{k_N} \left( \text{Tr}_{A_{k_N}^O} \Omega_\theta(h) \right) W_{(\mathcal{N} \setminus k_N, k_N)} \right) \\ & \text{s.t.} \begin{cases} \sum_{k_1} \text{Tr} W_{(\emptyset, k_1)} = 1, \\ \forall n = 1, \dots, N-1, \forall \mathcal{K}_n, \\ \sum_{k_{n+1} \in \mathcal{N} \setminus \mathcal{K}_n} \text{Tr}_{A_{k_{n+1}}^I} W_{(\mathcal{K}_n, k_{n+1})} \\ \quad = \sum_{k_n \in \mathcal{K}_n} W_{(\mathcal{K}_n \setminus k_n, k_n)} \otimes \mathbb{1}_{A_{k_n}^O}, \\ \forall n = 1, \dots, N, \forall \mathcal{K}_{n-1}, k_n, W_{(\mathcal{K}_{n-1}, k_n)} \geq 0. \end{cases} \end{aligned} \quad (\text{B29})$
$\begin{aligned} \mathcal{W} &= \bigoplus_{n=1}^N \bigoplus_{(k_1, \dots, k_n)} W_{(k_1, \dots, k_n)}, \\ \mathcal{A} &= \left( \bigoplus_{n=1}^{N-1} \bigoplus_{(k_1, \dots, k_n)} \mathbb{0}_{A_{\{k_1, \dots, k_{n-1}\}}^{IO} A_{k_n}^I} \right) \oplus \bigoplus_{(k_1, \dots, k_N)} \text{Tr}_{A_{k_N}^O} \Omega_\theta(h), \\ \Phi(\mathcal{W}) &= \bigoplus_{n=0}^{N-1} \bigoplus_{(k_1, \dots, k_n)} \Phi_{(k_1, \dots, k_n)}(\mathcal{W}) \\ & \text{with} \begin{cases} \Phi_\emptyset(\mathcal{W}) = \sum_{k_1} \text{Tr} W_{(k_1)}, \\ \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ \Phi_{(k_1, \dots, k_n)}(\mathcal{W}) = \sum_{k_{n+1}} \text{Tr}_{A_{k_{n+1}}^I} W_{(k_1, \dots, k_n, k_{n+1})} \\ \quad - W_{(k_1, \dots, k_n)} \otimes \mathbb{1}_{A_{k_n}^O}, \end{cases} \\ \mathcal{B} &= \mathbb{1} \oplus \bigoplus_{n=1}^{N-1} \bigoplus_{(k_1, \dots, k_n)} \mathbb{0}_{A_{\{k_1, \dots, k_n\}}^{IO}}. \end{aligned} \quad (\text{B30})$	$\begin{aligned} \mathcal{W} &= \bigoplus_{n=1}^N \bigoplus_{\mathcal{K}_{n-1}, k_n} W_{\mathcal{K}_{n-1}, k_n}, \\ \mathcal{A} &= \left( \bigoplus_{n=1}^{N-1} \bigoplus_{\mathcal{K}_{n-1}, k_n} \mathbb{0}_{A_{\mathcal{K}_{n-1}}^{IO} A_{k_n}^I} \right) \oplus \bigoplus_{k_N} \text{Tr}_{A_{k_N}^O} \Omega_\theta(h), \\ \Phi(\mathcal{W}) &= \bigoplus_{n=0}^{N-1} \bigoplus_{\mathcal{K}_n} \Phi_{\mathcal{K}_n}(\mathcal{W}) \\ & \text{with} \begin{cases} \Phi_\emptyset(\mathcal{W}) = \sum_{k_1} \text{Tr} W_{(\emptyset, k_1)}, \\ \forall n = 1, \dots, N-1, \forall \mathcal{K}_n, \\ \Phi_{\mathcal{K}_n}(\mathcal{W}) = \sum_{k_{n+1} \in \mathcal{N} \setminus \mathcal{K}_n} \text{Tr}_{A_{k_{n+1}}^I} W_{(\mathcal{K}_n, k_{n+1})} \\ \quad - \sum_{k_n \in \mathcal{K}_n} W_{(\mathcal{K}_n \setminus k_n, k_n)} \otimes \mathbb{1}_{A_{k_n}^O}, \end{cases} \\ \mathcal{B} &= \mathbb{1} \oplus \bigoplus_{n=1}^{N-1} \bigoplus_{\mathcal{K}_n} \mathbb{0}_{A_{\mathcal{K}_n}^{IO}}. \end{aligned} \quad (\text{B31})$
$\mathcal{S} = \bigoplus_{n=0}^{N-1} \bigoplus_{(k_1, \dots, k_n)} S_{(k_1, \dots, k_n)} \in \bigoplus_{n=0}^{N-1} \bigoplus_{(k_1, \dots, k_n)} \mathcal{L}(\mathcal{H}_{A_{\{k_1, \dots, k_n\}}^{IO}}). \quad (\text{B32})$	$\mathcal{S} = \bigoplus_{n=0}^{N-1} \bigoplus_{\mathcal{K}_n} S_{\mathcal{K}_n} \in \bigoplus_{n=0}^{N-1} \bigoplus_{\mathcal{K}_n} \mathcal{L}(\mathcal{H}_{A_{\mathcal{K}_n}^{IO}}). \quad (\text{B33})$
$\begin{aligned} \Phi^\dagger(\mathcal{S}) &= \bigoplus_{n=1}^N \bigoplus_{(k_1, \dots, k_n)} (\Phi^\dagger)_{(k_1, \dots, k_n)}(\mathcal{S}) \\ & \text{with} \begin{cases} \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ (\Phi^\dagger)_{(k_1, \dots, k_n)}(\mathcal{S}) = S_{(k_1, \dots, k_{n-1})} \otimes \mathbb{1}_{A_{k_n}^I} \\ \quad - \text{Tr}_{A_{k_n}^O} S_{(k_1, \dots, k_n)}, \\ \forall (k_1, \dots, k_N), \\ (\Phi^\dagger)_{(k_1, \dots, k_N)}(\mathcal{S}) = S_{(k_1, \dots, k_{N-1})} \otimes \mathbb{1}_{A_{k_N}^I}. \end{cases} \end{aligned} \quad (\text{B34})$	$\begin{aligned} \Phi^\dagger(\mathcal{S}) &= \bigoplus_{n=1}^N \bigoplus_{\mathcal{K}_{n-1}, k_n} (\Phi^\dagger)_{(\mathcal{K}_{n-1}, k_n)}(\mathcal{S}) \\ & \text{with} \begin{cases} \forall n = 1, \dots, N-1, \forall \mathcal{K}_n, \forall k_n \in \mathcal{K}_n, \\ (\Phi^\dagger)_{(\mathcal{K}_n \setminus k_n, k_n)}(\mathcal{S}) = S_{\mathcal{K}_n \setminus k_n} \otimes \mathbb{1}_{A_{k_n}^I} \\ \quad - \text{Tr}_{A_{k_n}^O} S_{\mathcal{K}_n}, \\ \forall k_N, (\Phi^\dagger)_{(\mathcal{N} \setminus k_N, k_N)}(\mathcal{S}) = S_{\mathcal{N} \setminus k_N} \otimes \mathbb{1}_{A_{k_N}^I}. \end{cases} \end{aligned} \quad (\text{B35})$

QC-CC (continued)	QC-QC (continued)
<p>Dual problem:</p> $\min_{\mathcal{S}} S_{\theta}$ $\text{s.t.} \begin{cases} \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ S_{(k_1, \dots, k_{n-1})} \otimes \mathbb{1}^{A_{k_n}^I} \geq \text{Tr}_{A_{k_n}^O} S_{(k_1, \dots, k_n)}, \\ \forall (k_1, \dots, k_N), \\ S_{(k_1, \dots, k_{N-1})} \otimes \mathbb{1}^{A_{k_N}^I} \geq \text{Tr}_{A_{k_N}^O} \Omega_{\theta}(h). \end{cases} \quad (\text{B36})$ <p>To verify Slater's condition for strong duality, one can consider <math>\mathcal{W} = \bigoplus_{n=1}^N \bigoplus_{(k_1, \dots, k_{n-1})} \frac{(N-n)!}{N!} \frac{\mathbb{1}^{A_{\{k_1, \dots, k_{n-1}\}}^O} \mathbb{1}^{A_{k_n}^I}}{d_{A_{\{k_1, \dots, k_n\}}^I}}</math>, which satisfies <math>\Phi(\mathcal{W}) = \mathcal{B}</math> and <math>\mathcal{W} &gt; 0</math>. (Finding a Hermitian <math>\mathcal{S}</math> such that <math>\Phi^{\dagger}(\mathcal{S}) \geq \mathcal{A}</math> is rather trivial.)</p> <p>Furthermore, a very similar construction to that in Eq. (B25) for the QC-FO case (essentially, replacing <math>S_n^{(l)}</math> by <math>S_{(k_1, \dots, k_n)}^{(l)}</math> and taking the identity operators in the appropriate spaces) then shows that the first inequalities above can without loss of generality be replaced by equalities, <math>S_{(k_1, \dots, k_{n-1})} \otimes \mathbb{1}^{A_{k_n}^I} = \text{Tr}_{A_{k_n}^O} S_{(k_1, \dots, k_n)}</math>.</p>	<p>Dual problem:</p> $\min_{\mathcal{S}} S_{\theta}$ $\text{s.t.} \begin{cases} \forall n = 1, \dots, N-1, \forall \mathcal{K}_n, \forall k_n \in \mathcal{K}_n, \\ S_{\mathcal{K}_n \setminus k_n} \otimes \mathbb{1}^{A_{k_n}^I} \geq \text{Tr}_{A_{k_n}^O} S_{\mathcal{K}_n}, \\ \forall k_N, S_{\mathcal{N} \setminus k_N} \otimes \mathbb{1}^{A_{k_N}^I} \geq \text{Tr}_{A_{k_N}^O} \Omega_{\theta}(h). \end{cases} \quad (\text{B37})$ <p>To verify Slater's condition for strong duality, one can consider <math>\mathcal{W} = \bigoplus_{n=1}^N \bigoplus_{\mathcal{K}_{n-1}, k_n} \frac{1}{n \binom{N}{n}} \frac{\mathbb{1}^{A_{\mathcal{K}_{n-1}}^O} \mathbb{1}^{A_{k_n}^I}}{d_{A_{\mathcal{K}_{n-1} \cup \{k_n\}}^I}}</math>, which satisfies <math>\Phi(\mathcal{W}) = \mathcal{B}</math> and <math>\mathcal{W} &gt; 0</math>. (Finding a Hermitian <math>\mathcal{S}</math> such that <math>\Phi^{\dagger}(\mathcal{S}) \geq \mathcal{A}</math> is rather trivial.)</p> <p>(Note that the construction of Eq. (B25) does not work here: except for the case <math>n = 1</math>, the inequalities above can in general not be replaced by equalities.)</p>
<p><math>S_{(k_1, \dots, k_{N-1})} \otimes \mathbb{1}^{A_{k_N}^I} \geq \text{Tr}_{A_{k_N}^O} \Omega_{\theta}(h)</math> in Eq. (B36) above can be linearized into</p> $\mathbf{A}_{\text{Schur}, k_N}(S_{(k_1, \dots, k_{N-1})}, h) \geq 0. \quad (\text{B38})$	<p><math>S_{\mathcal{N} \setminus k_N} \otimes \mathbb{1}^{A_{k_N}^I} \geq \text{Tr}_{A_{k_N}^O} \Omega_{\theta}(h)</math> in Eq. (B37) above can be linearized into</p> $\mathbf{A}_{\text{Schur}, k_N}(S_{\mathcal{N} \setminus k_N}, h) \geq 0. \quad (\text{B39})$
<p>Optimal QFI:</p> $J_{\theta}^{\text{QC-CC}}(C_{\theta}^{\otimes N})$ $= \min_{h, \mathcal{S}} S_{\theta}$ $\text{s.t.} \begin{cases} \forall n = 1, \dots, N-1, \forall (k_1, \dots, k_n), \\ S_{(k_1, \dots, k_{n-1})} \otimes \mathbb{1}^{A_{k_n}^I} = \text{Tr}_{A_{k_n}^O} S_{(k_1, \dots, k_n)}, \\ \forall (k_1, \dots, k_N), \mathbf{A}_{\text{Schur}, k_N}(S_{(k_1, \dots, k_{N-1})}, h) \geq 0. \end{cases} \quad (\text{B40})$	<p>Optimal QFI:</p> $J_{\theta}^{\text{QC-QC}}(C_{\theta}^{\otimes N})$ $= \min_{h, \mathcal{S}} S_{\theta}$ $\text{s.t.} \begin{cases} \forall n = 1, \dots, N-1, \forall \mathcal{K}_n, \forall k_n \in \mathcal{K}_n, \\ S_{\mathcal{K}_n \setminus k_n} \otimes \mathbb{1}^{A_{k_n}^I} \geq \text{Tr}_{A_{k_n}^O} S_{\mathcal{K}_n}, \\ \forall k_N, \mathbf{A}_{\text{Schur}, k_N}(S_{\mathcal{N} \setminus k_N}, h) \geq 0. \end{cases} \quad (\text{B41})$
<p>The optimal process <math>\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{QC-CC}</math> is recovered by solving Eq. (B28) after fixing <math>h = h^{(\text{opt})}</math> obtained from Eq. (B40) and adding the constraint that <math>\mathbf{C}_{\theta}^{0\dagger} \tilde{W}^T (\dot{\mathbf{C}}_{\theta}^0 - i \mathbf{C}_{\theta}^0 h^{(\text{opt})}) \in \mathbb{H}_q</math>, as in Eq. (B12)—and with <math>\tilde{W} = \sum_{(k_1, \dots, k_N)} W_{(N)} \otimes \mathbb{1}^{A_N^O}</math> (see Eq. (23)). Recall here that the purification <math>W^{(\text{opt})} =  w\rangle\langle w </math> of <math>\tilde{W}^{(\text{opt})}</math> is however not guaranteed to be in QC-CC.</p>	<p>The optimal process <math>\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{QC-QC}</math> is recovered by solving Eq. (B29) after fixing <math>h = h^{(\text{opt})}</math> obtained from Eq. (B41) and adding the constraint that <math>\mathbf{C}_{\theta}^{0\dagger} \tilde{W}^T (\dot{\mathbf{C}}_{\theta}^0 - i \mathbf{C}_{\theta}^0 h^{(\text{opt})}) \in \mathbb{H}_q</math>, as in Eq. (B12)—and with <math>\tilde{W} = \sum_{k_N \in \mathcal{N}} W_{(\mathcal{N} \setminus k_N, k_N)} \otimes \mathbb{1}^{A_{k_N}^O}</math> (see Eq. (24)). The optimal process <math>W^{(\text{opt})} =  w\rangle\langle w  \in \text{QC-QC}</math> is then obtained by purifying <math>\tilde{W}^{(\text{opt})}</math>.</p>

### c. SDP problems for the QC-Par, QC-Sup and Gen classes

For completeness, let us briefly write the explicit SDP problems one needs to consider for the remaining classes considered in this paper. These can be obtained in a similar fashion to the previous ones, or as in Ref. [19] (up to the difference mentioned in Footnote 19).

For the QC-Par class:

$$J_{\theta}^{\text{QC-Par}}(C_{\theta}^{\otimes N}) = \min_{h, S_0} S_0$$

$$\text{s.t. } S_0 \mathbb{1}^{A_{\mathcal{N}}} \geq \text{Tr}_{A_{\mathcal{N}}^O} \Omega_{\theta}(h), \quad (\text{B42})$$

with  $S_0 \in \mathbb{R}$ , and where the constraint can be linearized into

$$\mathbf{A}_{\text{Schur}}(S_0 \mathbb{1}^{A_{\mathcal{N}}}, \mathbf{C}_{\mathcal{N}}(h)) \geq 0 \quad (\text{B43})$$



with (similarly to Eq. (B17), with clear enough notation)

$$C_N(h) = 2 \left( \langle j_1, \dots, j_N | A_1^O \dots A_N^O (\dot{C}_\theta^0 - i C_\theta^0 h)^* \rangle_{j_1, \dots, j_N} \right). \quad (\text{B44})$$

The optimal  $\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{QC-Par}$  (and  $W^{(\text{opt})} = |w\rangle\langle w| \in \text{QC-Par}$ , after purification) is recovered by solving Eq. (B12), translating the constraint  $\tilde{W} \in \text{Tr}_F \text{QC-Par}$  into (see Eq. (17))

$$\begin{cases} \tilde{W} = W_{(I)} \otimes \mathbb{1}^{A_N^O}, \\ \text{Tr } W_{(I)} = 1, \\ W_{(I)} \geq 0. \end{cases} \quad (\text{B45})$$

For the QC-Sup class:

$$\begin{aligned} & J_\theta^{\text{QC-Sup}}(C_\theta^{\otimes N}) \\ & = \min_{h, S} S_0 \\ \text{s.t. } \forall \pi, & \begin{cases} S_0 \mathbb{1}^{A_{\pi(1)}} = \text{Tr}_{A_{\pi(1)}^O} S_{(\pi,1)}, \\ \forall n = 2, \dots, N-1, \\ S_{(\pi, n-1)} \otimes \mathbb{1}^{A_{\pi(n)}} = \text{Tr}_{A_{\pi(n)}^O} S_{(\pi, n)}, \\ S_{(\pi, N-1)} \otimes \mathbb{1}^{A_{\pi(N)}} \geq \text{Tr}_{A_{\pi(N)}^O} \Omega_\theta(h), \end{cases} \end{aligned} \quad (\text{B46})$$

where the variables in the optimization (all encapsulated in  $S$ , as done previously) are taken to be  $S_0 \in \mathbb{R}$  and  $S_{(\pi, n)} \in \mathcal{L}(\mathcal{H}^{A_{\{\pi(1), \dots, \pi(n)\}}^O})$  for all  $\pi$  and  $n = 1, \dots, N-1$ ; and where the last constraints (for each  $\pi$ ) can be linearized into

$$\mathbf{A}_{\text{Schur}, \pi(N)}(S_{(\pi, N-1)}, h) \geq 0. \quad (\text{B47})$$

The optimal  $\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{QC-Sup}$  (and  $W^{(\text{opt})} = |w\rangle\langle w| \in \text{QC-Sup}$ , after purification) is recovered by solving Eq. (B12), translating the constraint  $\tilde{W} \in \text{Tr}_F \text{QC-Sup}$  into (see Eq. (20), replacing  $q_\pi W_\pi$  by  $\tilde{W}_\pi$ )

$$\begin{cases} \tilde{W} = \sum_\pi \tilde{W}_\pi, \\ \forall \pi, \tilde{W}_\pi \in \text{Tr}_F(\text{QC-FO}_\pi^\times), \\ \text{Tr } \tilde{W} = d_{A_N^O}, \end{cases} \quad (\text{B48})$$

where  $\text{QC-FO}_\pi^\times$  denotes the class of process matrices that are proportional (with a nonnegative proportionality factor) to matrices in  $\text{QC-FO}_\pi$ , or more explicitly,

$$\begin{cases} \tilde{W} = \sum_\pi W_{(\pi, N)} \otimes \mathbb{1}^{A_{\pi(N)}^O}, \\ \sum_\pi \text{Tr } W_{(\pi, 1)} = 1, \\ \forall \pi, \forall n = 1, \dots, N-1, \\ \text{Tr}_{A_{\pi(n+1)}^O} W_{(\pi, n+1)} = W_{(\pi, n)} \otimes \mathbb{1}^{A_{\pi(n)}^O}, \\ \forall \pi, \forall n = 1, \dots, N, W_{(\pi, n)} \geq 0 \end{cases} \quad (\text{B49})$$

(with each  $W_{(\pi, n)} \in \mathcal{L}(\mathcal{H}^{A_{\{\pi(1), \dots, \pi(n-1)\}}^O} A_{\pi(n)}^O)$ ).

For the Gen class, finally:

$$\begin{aligned} & J_\theta^{\text{Gen}}(C_\theta^{\otimes N}) \\ & = \min_{h, S} \text{Tr}[S] / d_{A_N^I} \\ \text{s.t. } & \begin{cases} \forall k = 1, \dots, N, [1 - A_k^I] \text{Tr}_{A_k^O} S = 0, \\ S \geq \Omega_\theta(h), \end{cases} \end{aligned} \quad (\text{B50})$$

with  $S \in \mathcal{L}(\mathcal{H}^{A_N^O})$ , and where the last constraint can be linearized into

$$\mathbf{A}_{\text{Schur}}(S, 2(\dot{C}_\theta^0 - i C_\theta^0 h)^*) \geq 0. \quad (\text{B51})$$

The optimal  $\tilde{W}^{(\text{opt})} \in \text{Tr}_F \text{Gen}$  (and  $W^{(\text{opt})} = |w\rangle\langle w| \in \text{Gen}$ , after purification) is recovered by solving Eq. (B12), translating the constraint  $\tilde{W} \in \text{Tr}_F \text{Gen}$  into (see Eq. (21))

$$\begin{cases} \forall \emptyset \subsetneq \mathcal{K} \subseteq \mathcal{N}, \prod_{k \in \mathcal{K}} [1 - A_k^O] (\text{Tr}_{A_{\mathcal{N} \setminus \mathcal{K}}^O} \tilde{W}) = 0, \\ \text{Tr } \tilde{W} = d_{A_N^O}, \\ \tilde{W} \geq 0. \end{cases} \quad (\text{B52})$$

### 3. On the importance of imposing Eq. (B10)

To wrap up this appendix, let us clarify the need for imposing Eq. (B10) when reconstructing the optimal process (in step B1h above).

Consider some continuous, real-valued function  $f(h, w)$ , defined over some compact domains for  $h$  and  $w$ . Fan's minimax theorem [48] says that the optimal values of the  $\min_h \max_w$  and  $\max_w \min_h$  optimization problems coincide:

$$\min_h \max_w f(h, w) = \max_w \min_h f(h, w). \quad (\text{B53})$$

However, the solutions to both problems are in general not the same: a solution to the  $\min_h \max_w$  is not necessarily a solution to the  $\max_w \min_h$  problem.

Suppose that  $(h^\downarrow, w^\downarrow)$  and  $(h^\uparrow, w^\uparrow)$  are solutions to the  $\min_h \max_w$  and  $\max_w \min_h$  problems, respectively. Then one clearly has

$$f(h^\downarrow, w^\downarrow) \geq f(h^\downarrow, w^\uparrow) \geq f(h^\uparrow, w^\uparrow), \quad (\text{B54})$$

and from Eq. (B53) it follows that both inequalities are in fact equalities [19], so that  $(h^\downarrow, w^\uparrow)$  is a solution to both the  $\min_h \max_w$  and  $\max_w \min_h$  problems: it corresponds to a saddle point.

Hence, once a solution  $(h^\downarrow, w^\downarrow)$  of the  $\min_h \max_w$  problem is known, we know there exists a solution to the  $\max_w \min_h$  problem with  $h = h^\downarrow = h^{(\text{opt})}$ . Simply optimizing  $w$  for this fixed value of  $h$ , however, is not enough, in general, to identify such a solution  $(h^{(\text{opt})}, w^{(\text{opt})})$  to the  $\max_w \min_h$  problem. Indeed, it could be that several values of  $w$ , other than the saddle point, give the same value of  $f(h^{(\text{opt})}, w) = f(h^{(\text{opt})}, w^{(\text{opt})})$ . Such a situation is illustrated in Fig. 8.

To make sure that one indeed finds a solution to the  $\max_w \min_h$  problem when optimizing  $w$  for the fixed  $h = h^{(\text{opt})}$ , one thus needs to also impose that for the optimal value

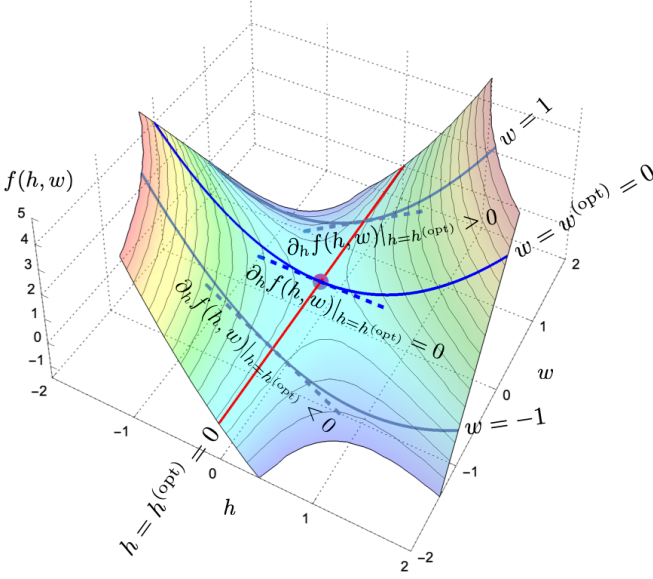


Figure 8. Plot of the function  $f(h, w) = h^2 + 2hw$ , as an illustration of a function that is convex in  $h$  and concave (even linear) in  $w$ .  $f(h, w)$  has a saddle point for  $(h^{(\text{opt})}, w^{(\text{opt})}) = (0, 0)$ . All points on the red line, for  $h = h^{(\text{opt})} = 0$ , are solutions to the  $\min_h \max_w f(h, w)$  problem. However, when fixing  $h = h^{(\text{opt})} = 0$  the function  $f(h^{(\text{opt})}, w)$  becomes constant in  $w$ : the saddle point, which is the only one that is also solution to the  $\max_w \min_h f(h, w)$  problem, can then only be found after imposing that the derivative of  $f(h, w)$  over  $h$  must be null, i.e.,  $\partial_h f(h, w)|_{h=h^{(\text{opt})}} = 0$ .

of  $w$ ,  $h^{(\text{opt})}$  is indeed a solution to  $\min_h f(w, h)$ . This can be done by requiring that  $\partial_h f(h, w)|_{h=h^{(\text{opt})}} = 0$  (which is a defining property for a saddle point). For the function  $f$  we consider in this work, this translates into Eq. (B10).

The situation just depicted for a general (convex-concave) function  $f(h, w)$  is indeed commonly encountered in the metrological context of our work. As a simple example, consider for instance the case of  $N = 1$  qubit depolarizing channel  $\mathcal{D}_\theta$  as in Eq. (9). Since  $N = 1$ , all the classes of processes considered in this work coincide. Taking  $\{|C_{\theta,j}^0\rangle\}_{j=1}^4$  to be an eigendecomposition of the Choi matrix of  $\mathcal{D}_\theta$  (with eigenvectors in the degenerate eigensubspace taken to be independent of  $\theta$ , except for their normalization) and following the approach described above, we find numerically, for all values of  $\theta$ , the optimal value of  $h$  to be  $h^{(\text{opt})} = 0$ .

Now, for this value one finds that  $\text{Tr}_{A^O} \Omega_\theta(h^{(\text{opt})}) = \frac{3}{(1-\theta)(1+3\theta)} \mathbb{1}^{A^I}$ . Since for all classes,  $\tilde{W}$  is of the form  $\tilde{W} = \rho^{A^I} \otimes \mathbb{1}^{A^O}$  with  $\text{tr}(\rho^{A^I}) = 1$ , one then finds that  $\text{Tr}(\Omega_\theta(h^{(\text{opt})}) \tilde{W}) = \frac{3}{(1-\theta)(1+3\theta)}$  is independent from  $\tilde{W}$ . While this indeed gives the correct value for a single qubit depolarizing channel QFI (see Eq. (31)), clearly one could not recover the optimal  $\tilde{W}$  (or  $\rho^{A^I}$ ) by directly optimizing  $\text{Tr}(\Omega_\theta(h^{(\text{opt})}) \tilde{W})$ . Imposing Eq. (B10) however, enforces  $\rho^{A^I}$  to be the maximally mixed state  $\frac{\mathbb{1}^{A^I}}{2}$ , which indeed (after purification into a maximally entangled 2-qubit state) gives the

optimal QFI of the depolarizing channel.<sup>23</sup>

### Appendix C: Numerical results for the QFI of the composed rotation & amplitude damping channels $\mathcal{A}_\theta$

We report in Tables I and II, and in Figures 9 and 10, the optimal values of the QFI obtained numerically (by solving the SDP) for the different classes, in the case of  $N = 2$  and  $N = 3$  queries (respectively) to the composed rotation & amplitude damping channel  $\mathcal{A}_\theta$  defined in Eq. (34). As mentioned in the main text, the QFI does not depend on the specific value of  $\theta$  for these channels. We recover the hierarchies obtained in Eqs. (35) and (36), respectively.

$u$	$J_\theta^{\text{QC-Par}}(A_\theta^{\otimes 2})$	$J_\theta^{\text{QC-FO}}(A_\theta^{\otimes 2}), J_\theta^{\text{QC-CC}}(A_\theta^{\otimes 2})$	$J_{\theta, \text{purif}}^{\text{QC-CC}}(A_\theta^{\otimes 2}), J_\theta^{\text{QC-Sup}}(A_\theta^{\otimes 2}), J_\theta^{\text{QC-QC}}(A_\theta^{\otimes 2}), J_\theta^{\text{Gen}}(A_\theta^{\otimes 2})$
0	4	4	4
0.1	3.5900	3.6477	3.6929
0.2	3.1605	3.2904	3.3709
0.3	2.7128	2.9273	3.0324
0.4	2.2500	2.5574	2.6760
0.5	1.7946	2.1793	2.2995
0.6	1.4314	1.7911	1.9005
0.7	1.1177	1.3900	1.4764
0.8	0.8117	0.9714	1.0240
0.9	0.4734	0.5260	0.5410
1	0	0	0

Table I. QFI for  $N = 2$  uses of the composed rotation & amplitude damping channel  $\mathcal{A}_\theta$ , for the different classes under consideration, for  $0 \leq u \leq 1$  (and any fixed value of  $\theta$ ). Notice that here (except for  $u = 0$  and  $u = 1$ )  $J_{\theta, \text{purif}}^{\text{QC-CC}}(A_\theta^{\otimes 2})$  is a strict upper bound on  $J_\theta^{\text{QC-CC}}(A_\theta^{\otimes 2})$ —which, in the case of  $N = 2$ , we know is equal to  $J_\theta^{\text{QC-FO}}(A_\theta^{\otimes 2})$ ; cf. Eq. (29).

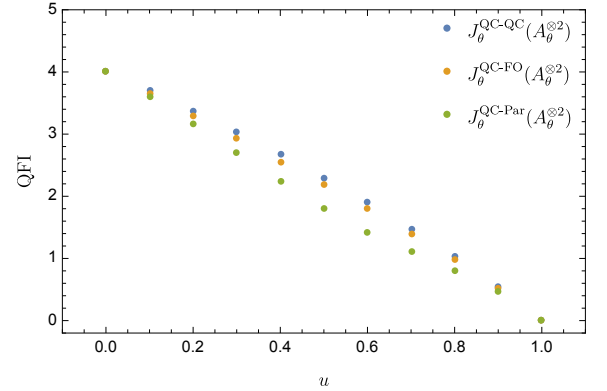


Figure 9. QFI for  $N = 2$  uses of the composed rotation & amplitude damping channel  $\mathcal{A}_\theta$ , as in Table I.

<sup>23</sup> Considering inputting half of a 2-qubit state  $|\psi_{\text{in}}\rangle = \alpha|0, 0\rangle + \beta|1, 1\rangle$  into  $\mathcal{D}_\theta$ , one finds (e.g., using Eq. (2)) the QFI of the 2-qubit output state to be  $\frac{1+3\theta+8|\alpha|^2|\beta|^2}{(1-\theta^2)(1+3\theta)}$ . This indeed clearly depends on the input state, and is optimal for a maximally entangled state.

$u$	$J_{\theta}^{\text{QC-Par}}(A_{\theta}^{\otimes 3})$	$J_{\theta}^{\text{QC-FO}}(A_{\theta}^{\otimes 3})$	$J_{\theta}^{\text{QC-CC}}(A_{\theta}^{\otimes 3}),$ $J_{\theta}^{\text{QC-Sup}}(A_{\theta}^{\otimes 3}),$ $J_{\theta}^{\text{QC-purif}}(A_{\theta}^{\otimes 3})$	$J_{\theta}^{\text{QC-QC}}(A_{\theta}^{\otimes 3})$	$J_{\theta}^{\text{Gen}}(A_{\theta}^{\otimes 3})$
0	9	9	9	9	9
0.1	7.6366	7.9238	8.1848	8.1986	8.2001
0.2	6.2628	6.8951	7.3643	7.3741	7.3751
0.3	4.9111	5.9140	6.5231	6.5242	6.5242
0.4	3.8614	4.9797	5.6417	5.6461	5.6468
0.5	3.1082	4.0905	4.7247	4.7373	4.7431
0.6	2.4344	3.2427	3.7859	3.8048	3.8150
0.7	1.8217	2.4300	2.8317	2.8564	2.8698
0.8	1.2786	1.6412	1.8710	1.8957	1.9086
0.9	0.7255	0.8553	0.9177	0.9260	0.9299
1	0	0	0	0	0

Table II. QFI for  $N = 3$  uses of the composed rotation & amplitude damping channel  $\mathcal{A}_{\theta}$ , for the different classes under consideration, for  $0 \leq u \leq 1$  (and any fixed value of  $\theta$ ). Recall that  $J_{\theta}^{\text{QC-CC}}(A_{\theta}^{\otimes 3})$  is only an upper bound on  $J_{\theta}^{\text{QC-CC}}(A_{\theta}^{\otimes 3})$ , which our approach does not allow us to access directly; cf. Eq. (30).

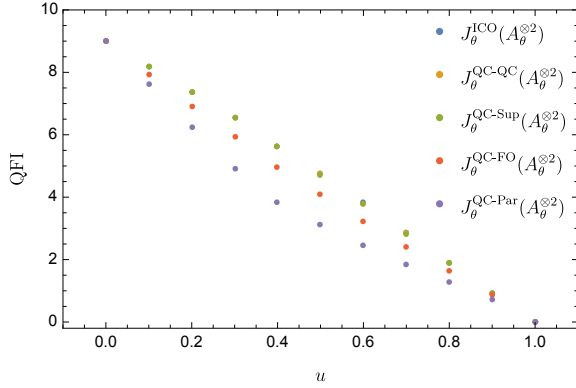


Figure 10. QFI for  $N = 3$  uses of the composed rotation & amplitude damping channel  $\mathcal{A}_{\theta}$ , as in Table II.

- 
- [1] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012), [arXiv:1105.4464 \[quant-ph\]](#).
- [2] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, *Phys. Rev. A* **88**, 022318 (2013), [arXiv:0912.0195 \[quant-ph\]](#).

- [3] L. Hardy, Probability theories with dynamic causal structure: a new framework for quantum gravity, [arXiv:gr-qc/0509120](#) (2005).
- [4] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, Bell's theorem for temporal order, *Nat. Commun.* **10**, 3772 (2019), [arXiv:1708.00248 \[quant-ph\]](#).
- [5] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, *Phys. Rev. A* **86**, 040301(R) (2012), [arXiv:1109.5154 \[quant-ph\]](#).
- [6] D. Ebler, S. Salek, and G. Chiribella, Enhanced communication with the assistance of indefinite causal order, *Phys. Rev. Lett.* **120**, 120502 (2018), [arXiv:1711.10165 \[quant-ph\]](#).
- [7] A. Feix, M. Araújo, and Č. Brukner, Quantum superposition of the order of parties as a communication resource, *Phys. Rev. A* **92**, 052326 (2015), [arXiv:1508.07840 \[quant-ph\]](#).
- [8] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, *New J. Phys.* **17**, 102001 (2015), [arXiv:1506.03776 \[quant-ph\]](#).
- [9] 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**,

<sup>24</sup> For instance, for  $N = 2$ , for the class of parallel strategies QC-Par and for  $u \lesssim 0.45$ , the SDP returns optimal processes of the form  $\tilde{W} = \text{Tr}_F W = (a|00\rangle\langle 00| + b|11\rangle\langle 11|)A_1^I A_2^I \otimes \mathbb{1}^{A_1^O} A_2^O$ , for some (nonnegative,  $u$ -dependent) values of  $a, b$ . These can be purified in the form  $W = |w\rangle\langle w|$  with  $|w\rangle = (\sqrt{a}|000\rangle + \sqrt{b}|111\rangle)A_1^I A_2^I F_1 \otimes |\mathbb{1}\rangle\rangle^{A_1^O} F_1 \otimes |\mathbb{1}\rangle\rangle^{A_2^O} F_2$ ; the QFI of the output state can then be calculated analytically, and optimized (for some fixed  $u$ ) over the values of  $a, b$ —which leads to  $J_{\theta}^{\text{QC-Par}}(A_{\theta}^{\otimes 2}) = \frac{4(1-u)^2}{(1-u/2)^2}$ . For  $u \gtrsim 0.45$  the optimal processes are found to be of the form  $\tilde{W} = \text{Tr}_F W = (a|00\rangle\langle 00| + b|11\rangle\langle 11| + c|01\rangle\langle 01| + c|10\rangle\langle 10|)A_1^I A_2^I \otimes \mathbb{1}^{A_1^O} A_2^O$ , rather, and expressing the corresponding optimal QFI analytically is already much more involved.

Recall that our SDPs allow us, for each class, to reconstruct the processes that give the optimal values of the QFI. Except in the simplest cases<sup>24</sup> it is, however, difficult to gain much insight or intuition from the form of the processes thus obtained numerically.

- 100502 (2016), [arXiv:1605.07372 \[quant-ph\]](#).
- [10] 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), [arXiv:1810.10457 \[quant-ph\]](#).
- [11] C. Mukhopadhyay, M. K. Gupta, and A. K. Pati, Superposition of causal order as a metrological resource for quantum thermometry, [arXiv:1812.07508 \[quant-ph\]](#) (2018).
- [12] M. Frey, Indefinite causal order aids quantum depolarizing channel identification, *Quantum Inf. Process.* **18**, 96 (2019).
- [13] X. Zhao, Y. Yang, and G. Chiribella, Quantum metrology with indefinite causal order, *Phys. Rev. Lett.* **124**, 190503 (2020), [arXiv:1912.02449 \[quant-ph\]](#).
- [14] M. Frey, Quantum network probing with indefinite routing, *Quantum Inf. Process.* **20**, 13 (2021).
- [15] A. Altherr and Y. Yang, Quantum metrology for non-Markovian processes, *Phys. Rev. Lett.* **127**, 060501 (2021), [arXiv:2103.02619 \[quant-ph\]](#).
- [16] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, *Phys. Rev. A* **103**, 032615 (2021), [arXiv:2104.06284 \[quant-ph\]](#).
- [17] F. Chapeau-Blondeau, Quantum parameter estimation on coherently superposed noisy channels, *Phys. Rev. A* **104**, 032214 (2021), [arXiv:2110.06715 \[quant-ph\]](#).
- [18] F. Chapeau-Blondeau, Indefinite causal order for quantum metrology with quantum thermal noise, *Phys. Lett. A* **447**, 128300 (2022), [arXiv:2211.01684 \[quant-ph\]](#).
- [19] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal strategies of quantum metrology with a strict hierarchy, *Phys. Rev. Lett.* **130**, 070803 (2023), [arXiv:2203.09758 \[quant-ph\]](#).
- [20] G. Chiribella and X. Zhao, Heisenberg-limited metrology with coherent control on the probes' configuration, [arXiv:2206.03052 \[quant-ph\]](#) (2022).
- [21] S. Kurdzialek, W. Gorecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, *Phys. Rev. Lett.* **131**, 090801 (2023), [arXiv:2212.08106 \[quant-ph\]](#).
- [22] A. Z. Goldberg, K. Heshami, and L. L. Sánchez-Soto, Evading noise in multiparameter quantum metrology with indefinite causal order, *Phys. Rev. Research* **5**, 033198 (2023), [arXiv:2309.07220 \[quant-ph\]](#).
- [23] F. Chapeau-Blondeau, Indefinite causal order for quantum phase estimation with pauli noise, *Fluct. Noise Lett.* **22**, 2350036 (2023), [arXiv:2312.02832 \[quant-ph\]](#).
- [24] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [25] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Generalized uncertainty relations: Theory, examples, and Lorentz invariance, *Annal. Phys.* **247**, 135 (1996), [arXiv:quant-ph/9507004](#).
- [26] 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, G. Chen, C.-F. Li, and G.-C. Guo, Experimental super-Heisenberg quantum metrology with indefinite gate order, [arXiv:2303.17223 \[quant-ph\]](#) (2023).
- [27] 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), [arXiv:1608.01683 \[quant-ph\]](#).
- [28] 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), [arXiv:1803.04302 \[quant-ph\]](#).
- [29] 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), [arXiv:1810.10238 \[quant-ph\]](#).
- [30] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Increasing communication capacity via superposition of order, *Phys. Rev. Research* **2**, 033292 (2020), [arXiv:1807.07383 \[quant-ph\]](#).
- [31] 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), [arXiv:1811.07526 \[quant-ph\]](#).
- [32] G. Rubino, G. Manzano, and Č. Brukner, Quantum superposition of thermodynamic evolutions with opposing time's arrows, *Commun. Phys.* **4**, 251 (2021), [arXiv:2008.02818 \[quant-ph\]](#).
- [33] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, *Phys. Rev. Research* **3**, 013093 (2021), [arXiv:2007.05005 \[quant-ph\]](#).
- [34] 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), [arXiv:2101.08796 \[quant-ph\]](#).
- [35] C. W. Helstrom, Minimum mean-squared error of estimates in quantum statistics, *Phys. Lett. A* **25**, 101 (1967).
- [36] C. W. Helstrom, Quantum detection and estimation theory, *J. Stat. Phys.* **1**, 231 (1969).
- [37] M. G. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inf.* **7**, 125 (2009), [arXiv:0804.2981 \[quant-ph\]](#).
- [38] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, *J. Phys. A: Math. Theor.* **41**, 255304 (2008).
- [39] Z. Huang, C. Macchiavello, and L. Maccone, Usefulness of entanglement-assisted quantum metrology, *Phys. Rev. A* **94**, 012101 (2016), [arXiv:1603.03993 \[quant-ph\]](#).
- [40] A. Fujiwara, Quantum channel identification problem, *Phys. Rev. A* **63**, 042304 (2001).
- [41] A. Fujiwara, Estimation of a generalized amplitude-damping channel, *Phys. Rev. A* **70**, 012317 (2004).
- [42] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, A purification postulate for quantum mechanics with indefinite causal order, *Quantum* **1**, 10 (2017), [arXiv:1611.08535 \[quant-ph\]](#).
- [43] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**, 285 (1975).
- [44] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, *Phys. Rev. Lett.* **101**, 060401 (2008), [arXiv:0712.1325 \[quant-ph\]](#).
- [45] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, *Phys. Rev. A* **80**, 022339 (2009), [arXiv:0904.4483 \[quant-ph\]](#).
- [46] M. Araújo, F. Costa, and Č. Brukner, Computational advantage from quantum-controlled ordering of gates, *Phys. Rev. Lett.* **113**, 250402 (2014), [arXiv:1401.8127 \[quant-ph\]](#).
- [47] G. Chiribella and D. Ebler, Optimal quantum networks and one-shot entropies, *New J. Phys.* **18**, 093053 (2016), [arXiv:1606.02394 \[quant-ph\]](#).
- [48] K. Fan, Minimax theorems, *Proc. Natl. Acad. Sci. USA* **39**, 42 (1953).
- [49] G. Gutoski and J. Watrous, Toward a general theory of quantum games, in *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, STOC '07 (2007) pp. 565–574, [arXiv:quant-ph/0611234](#).



- [50] J. Wechs, A. A. Abbott, and C. Branciard, On the definition and characterisation of multipartite causal (non)separability, *New J. Phys.* **21**, 013027 (2019), [arXiv:1807.10557 \[quant-ph\]](#).
- [51] T. Purves and A. J. Short, Quantum theory cannot violate a causal inequality, *Phys. Rev. Lett.* **127**, 110402 (2021), [arXiv:2101.09107 \[quant-ph\]](#).
- [52] A. Fujiwara and H. Imai, Quantum parameter estimation of a generalized Pauli channel, *J. Phys. A: Math. Gen.* **36**, 8093 (2003).
- [53] N. Johnston, A. Cosentino, and V. Russo, *QETLAB: QETLAB v0.9*, Zenodo (2016).
- [54] A. A. Abbott, M. Mhalla, and P. Pocreau, Quantum query complexity of Boolean functions under indefinite causal order, [arXiv:2307.10285 \[quant-ph\]](#) (2023).
- [55] R. Mothe, M. Grothus, A. A. Abbott, and C. Branciard, On correlations and quantum processes with dynamical causal orders, in preparation.
- [56] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, *Phys. Lett. A* **183**, 14 (1993).
- [57] J. Watrous, *The theory of quantum information* (Cambridge university press, 2018).
- [58] R. A. Horn and F. Zhang, Basic properties of the Schur complement, in *The Schur Complement and Its Applications*, edited by F. Zhang (Springer US, 2005).