Channel capacity enhancement with indefinite causal order

Nicolas Loizeau, Alexei Grinbaum

Channel capacity enhancement with indefinite causal order

Nicolas Loizeau

Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

Alexei Grinbaum

CEA-Saclay, IRFU/Larsim, 91191 Gif-sur-Yvette Cedex, France

(Dated: November 1, 2019)

Classical communication capacity of a channel can be enhanced either through a device called a 'quantum switch' or by putting the channel in a quantum superposition. The gains in the two cases, although different, have their origin in the use of a quantum resource, but is it the same resource? Here this question is explored through simulating large sets of random channels. We find that quantum superposition always provides an advantage, while the quantum switch does not: it can either increase or decrease communication capacity. The origin of this discrepancy can be attributed to a subtle combination of superposition and non-commutativity.

I. INTRODUCTION

In spacetime, events A and B can be in three causal relations: either A is before B, B is before A, or A and B are causally separated, i.e. they lie on a spacelike interval. Quantum mechanics admits causal structures that do not correspond to any of these cases. Heuristically, this can be pictured as putting the order between A and B in a quantum superposition. More precisely, several approaches to indefinite causal orders have been proposed using 'process matrix' or 'quantum switch' [1–6]. While these approaches are not strictly equivalent mathematically, all of them support one underlying idea: an indefinite causal order is an inherently quantum phenomenon that sheds new light on a notion hitherto explored mainly in spacetime theories. This phenomenon has recently been observed experimentally in several implementations of the quantum switch [7–12].

To gauge precisely the new element brought by quantum theory into the study of causality, quantum control on causal order can be treated as a resource that provides non-classical communication advantage, i.e., two noisy channels in a quantum switch can transmit more information than any of these channels individually [13]. This approach has the benefit of immediately clarifying the physical interest of the quantum switch, however it relies on a currently unresolved question whether any local party can operationally exercise such quantum control [14]. In this work we assume that a positive heuristic has been given by the empirical work: quantum control of the causal order via the quantum switch has been obtained experimentally. In what follows we strive to achieve a better theoretical understanding of the advantage demonstrated in such setups. In particular, a standing problem concerns the origin of this advantage: to deny that the quantum switch is an independent resource, it has been argued that the one-pass quantum superposition of two channels, without the indefinite causal order, already leads to a similar result [15, 16].

After introducing basic mathematical concepts in Section II, we explore the controversial origin of this non-classical advantage in Section III. To this end, we simulate large sets of random channels and use them to compare communication gains from the quantum switch and the one-pass superposition. In Section V, we argue that the advantage for the quantum switch has its origin in two separate factors. One is quantum superposition; the other is non-commutativity of the Kraus decompositions of the channels. A combination of these factors can be significantly more beneficial than the advantage gained from the superposition alone, but in other cases it can also be much less advantageous. When the indefinite causal order is realized through a quantum switch, the gain provided by this resource is essentially due to this combination.

II. MATHEMATICAL FRAMEWORK

A quantum system going through a quantum channel is modelled by a completely positive trace preserving linear map on its state Hilbert space $\mathcal{H}$. Any such map $\mathcal{C}$ can be represented by a set of Kraus operators $\{K_i\} \subset \mathcal{L}(\mathcal{H})$ such as [17–19]:

$$\mathcal{C}(\rho) = \sum_i K_i \rho K_i^\dagger \quad \text{and} \quad \sum_i K_i^\dagger K_i = \mathbb{1}. \qquad (1)$$

This decomposition is not unique: if $\mathcal{C}$ and $\mathcal{C}'$ have Kraus operators $\{K_i\}$ and $\{K_i'\}$ respectively, then $\mathcal{C}$ implements the same channel as $\mathcal{C}'$ if and only if there exists a unitary operator u such as:

$$K_i = \sum_j u_{ij} K_j'. \qquad (2)$$

A quantum switch $\mathcal{C}_0 \bowtie \mathcal{C}_1$ between channels $\mathcal{C}_0$ and $\mathcal{C}_1$ is a new channel that puts in a superposition two differently ordered compositions $\mathcal{C}_0 \circ \mathcal{C}_1$ and $\mathcal{C}_1 \circ \mathcal{C}_0$ (Figure 1). It acts on $\mathcal{L}(\mathcal{H}^c) \otimes \mathcal{L}(\mathcal{H}^t)$, where c stands for control and t for target. Unless stated otherwise, both of these Hilbert spaces are taken to be two-dimensional. The switch $\mathcal{C}_0 \bowtie \mathcal{C}_1$ is a higher-order operation defined through its Kraus decomposition

$$V_{ij} = |0\rangle\langle 0|^c \otimes (K_i^1 K_j^0)^t + |1\rangle\langle 1|^c \otimes (K_j^0 K_i^1)^t, \quad (3)$$

where K_i^0 and K_j^1 are the Kraus operators of $\mathcal{C}_0$ and $\mathcal{C}_1$ respectively. To get a heuristic picture, one may think about the target as a subsystem that passes through the channels, while the control subsystem in a generic state $|\psi\rangle^c = \alpha |0\rangle^c + \beta |1\rangle^c, |\alpha|^2 + |\beta|^2 = 1$, determines the order of passage. This is easy to comprehend in the case of unitary $U_0, U_1 \in \mathcal{L}(\mathcal{H}^t)$ with the subsystems in pure states:

$$(U_0 \bowtie U_1) |\psi\rangle^c \otimes |\psi\rangle^t = \alpha |0\rangle^c \otimes U_0 U_1 |\psi\rangle^t \quad (4)$$
$$+ \beta |1\rangle^c \otimes U_1 U_0 |\psi\rangle^t.$$

One can see that, depending on the state of the control qubit, the order in which the target undergoes different operations is *switched*. If, for example, the state of the control qubit is $|+\rangle^c = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)^c$, then the quantum switch yields a balanced superposition of two orders.
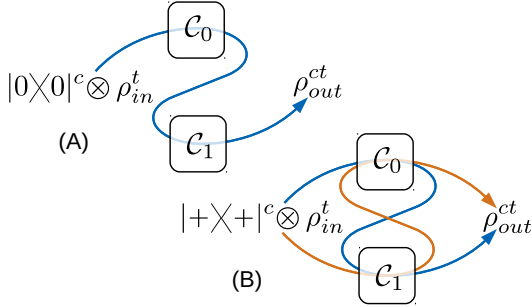

FIG. 1. Quantum circuit representing the quantum switch. Target system ρ_{in}^t passes through $\mathcal{C}_0 \circ \mathcal{C}_1$ and $\mathcal{C}_1 \circ \mathcal{C}_0$ in a superposition determined by control qubit c. (A): the control is in the reduced state $|0\rangle\langle 0|^c$ and the target passes only through $\mathcal{C}_0 \circ \mathcal{C}_1$. (B): the control is in the reduced state $|+\rangle\langle +|^c$ and the target passes through a balanced superposition of $\mathcal{C}_0 \circ \mathcal{C}_1$ and $\mathcal{C}_1 \circ \mathcal{C}_0$.

Somewhat paradoxically, classical information can be transmitted though the quantum switch between two totally depolarizing channels. The Holevo capacity of a channel $\chi(\mathcal{C})$ is defined as $\chi(\mathcal{C}) = \max_{\{p_a, \rho_a\}} I(A;B)_\nu$, where $\{\rho_a\}$ are the possible inputs of the channel with probabilities p_a and $I(A;B)_\nu$ is the quantum mutual information calculated on the state $\nu = \sum_a p_a |a\rangle\langle a|_A \otimes \mathcal{C}(\rho_a)_B$ [20]. In terms of the von Neumann entropy H, if $\rho = \sum_a p_a \rho_a$, then

$$\chi(\mathcal{C}) = \max_{\{p_a, \rho_a\}} \left[H(\mathcal{C}(\rho)) - \sum_a p_a H(\mathcal{C}(\rho_a)) \right]. \quad (5)$$

This maximum can be reached on no more than $N^2 = 4$ pure states [21], where N is the dimension of the target subsystem.

The Holevo capacity of the quantum switch between two totally depolarizing channels acting on a qubit equals

$\chi = -\frac{8}{3} - \frac{5}{8} \log_2 \frac{5}{8} \approx 0.05$ [13]. However, a similar advantage occurs if the channels are put in a superposition $C_0 \boxplus C_1$ (Figure 2). The one-pass superposition of $\mathcal{C}_0$ and $\mathcal{C}_1$ is defined through its Kraus operators

$$W_{ij} = \frac{1}{2} |0\rangle\langle 0|^c \otimes K_i^{0t} + \frac{1}{2} |1\rangle\langle 1|^c \otimes K_j^{1t}. \quad (6)$$

As before, in the case of unitary U_0 and U_1 the definition can be simplified:

$$(U_0 \boxplus U_1) |\psi\rangle^c \otimes |\psi\rangle^t = \alpha |0\rangle^c \otimes U_0 |\psi\rangle^t + \beta |1\rangle^c \otimes U_1 |\psi\rangle^t. \quad (7)$$
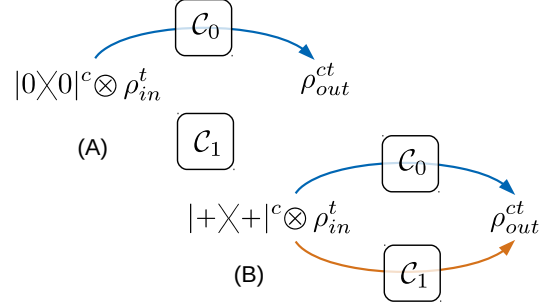

FIG. 2. Quantum circuit representing the one-pass superposition of $\mathcal{C}_0$ and $\mathcal{C}_1$. Target system ρ_{in}^t passes through $\mathcal{C}_0$ and/or $\mathcal{C}_1$ in a superposition determined by control qubit ρ^c. (A): the control is in the reduced state $|0\rangle\langle 0|^c$ and the target passes only through $\mathcal{C}_0$. (B): the control is in the reduced state $|+\rangle\langle +|^c$ and the target passes through a balanced superposition of $\mathcal{C}_0$ and $\mathcal{C}_1$.

Starting from a well-known result in quantum communication [22], it has recently been shown that $\mathcal{C}_0 \boxplus \mathcal{C}_1$ has a greater Holevo capacity than $\mathcal{C}_0 \bowtie \mathcal{C}_1$ if $\mathcal{C}_0$ and $\mathcal{C}_1$ are totally depolarizing. A lower bound is $\chi(\mathcal{C}_0 \boxplus \mathcal{C}_1) \geq 0.16$ [15].

III. RESULTS

The indefinite causal order provides an indisputable advantage in terms of Holevo capacity but this advantage is not systematic. To explore this situation, we randomly generate pairs of quantum channels $\mathcal{C}_0$ and $\mathcal{C}_1$ and numerically compute Holevo capacities of $\mathcal{C}_0 \bowtie \mathcal{C}_1$ and $\mathcal{C}_0 \boxplus \mathcal{C}_1$ for a control qubit in the reduced state $|+\rangle\langle +|^c$. Figure 3, generated on two sets of 1000 channels each, shows the absence of any obvious correlation between $\chi(\mathcal{C}_0 \bowtie \mathcal{C}_1)$ and $\chi(\mathcal{C}_0 \boxplus \mathcal{C}_1)$. After three such runs, the average ratio $\chi(\mathcal{C} \bowtie \mathcal{C})/\chi(\mathcal{C}_0 \boxplus \mathcal{C}_1)$ is stable around 0.9, meaning that on average the one-pass superposition gives a slightly better advantage than the quantum switch. However, individual channels can exhibit vastly different behaviour.

To study the combination of a channel with itself, we set $\mathcal{C}_0 = \mathcal{C}_1 = \mathcal{C}$. Figure 4 shows that $\chi(\mathcal{C} \boxplus \mathcal{C})$ varies significantly as one applies (2) to change the Kraus decomposition of a channel, while $\chi(\mathcal{C} \bowtie \mathcal{C})$ remains fixed.
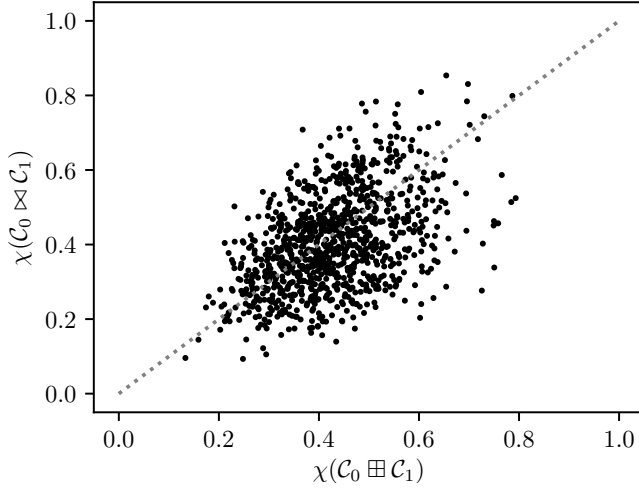
FIG. 3. Comparison of Holevo capacities of the quantum switch and the one-pass superposition. Each point corresponds to a couple of randomly generated quantum channels. Holevo capacities are estimated numerically.
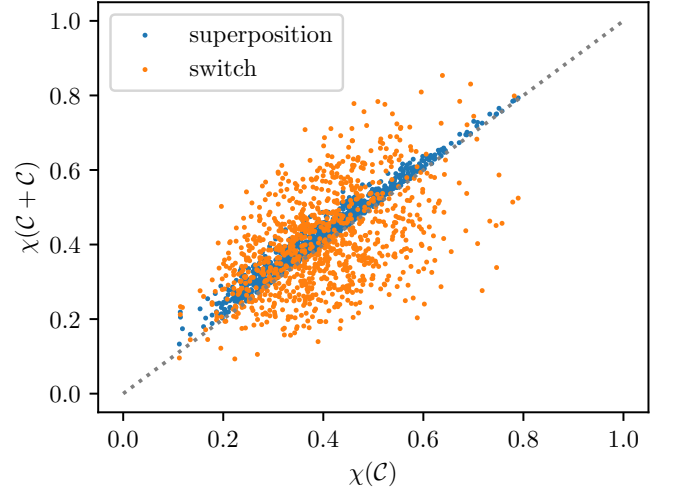

FIG. 5. Comparison of Holevo capacities of self-switch and self-superposition of 1000 random channels. Holevo capacity always increases in the latter case but no such regularity exists in the former. The $+$ sign stands alternatively for $\bowtie$ or $\boxplus$.
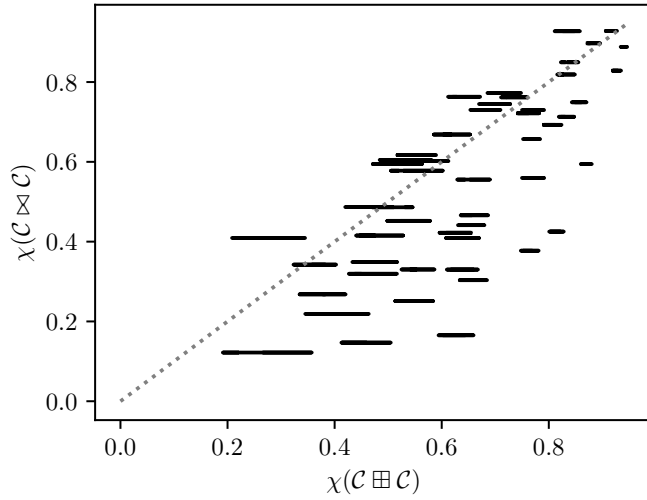

FIG. 4. Comparison of Holevo capacities of the quantum switch and the one-pass superposition for different implementations of 50 randomly generated channels. Each line corresponds to 200 different implementations of channel $\mathcal{C}$ due to the freedom in the Kraus decomposition. $\chi(\mathcal{C} \boxplus \mathcal{C})$ depends on the implementation, while $\chi(\mathcal{C} \bowtie \mathcal{C})$ is independent of the implementation.

FIG. 6. Comparison of Holevo capacities of the quantum switch and the one-pass superposition of a random channel $\mathcal{C}$ combined with a totally depolarizing channel $\mathcal{N}$. Each point corresponds to a random channel $\mathcal{C}$. Dotted line are linear functions with the slopes 1 and $1/2$. The $+$ sign stands alternatively for $\bowtie$ or $\boxplus$.

Figure 5 provides a comparison between the Holevo capacities of self-switch and self-superposition. The latter always increases channel capacity: $\chi(\mathcal{C}) < \chi(\mathcal{C} \boxplus \mathcal{C})$ as we prove below, while this is not true for the former. If the number of subsystems is increased, e.g. in the case of the 3-switch [12], the picture remains very similar to the one shown here. To make sure that this effect is not only due to self-switching, we explore the Holevo ca-
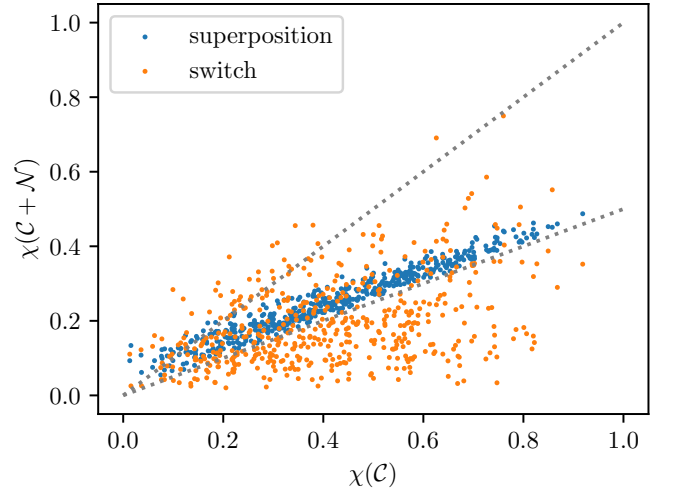
pacity of a composition between a random channel and a totally depolarizing channel $\mathcal{N}$ with Kraus operators $\{\frac{\mathbb{1}}{2}, \frac{\sigma_x}{2}, \frac{\sigma_y}{2}, \frac{\sigma_z}{2}\}$ (Fig. 6). All generated channels verify $\chi(\mathcal{C})/2 < \chi(\mathcal{C} \boxplus \mathcal{N})$, whereas this inequality does not hold for the quantum switch.

IV. METHODS

A random channel $\mathcal{C}$ is obtained by generating a random set of Kraus operators. To get the latter, we generate a random Choi matrix. Let $X \in \mathcal{L}(\mathcal{H})_A \otimes \mathcal{L}(\mathcal{H})_B$ be a Ginibre complex random matrix and $Y = \mathrm{Tr}_A XX^\dagger$. Then the following matrix is random Choi [23]:

$$C = \left(\mathbb{1}_A \otimes \frac{1}{\sqrt{Y}}\right) XX^\dagger \left(\mathbb{1}_A \otimes \frac{1}{\sqrt{Y}}\right). \qquad (8)$$

Holevo capacity $\chi(\mathcal{C})$ is computed by solving the optimization problem (5) using the basin-hopping method [24]. States ρ_a are parameterized as $\rho_a = |a\rangle\langle a|$, $|a\rangle = (cos(\theta_a), sin(\theta_a)e^{i\phi_a})^T$, $(\phi_a, \theta_a) \in [0,\pi] \times [0,2\pi]$. Each step of the basin-hopping method consists of: a) random perturbation of the parameters, b) local minimization using the BFGS method, c) acceptance or rejection of the new parameters based on the minimized function value and the metropolis test [25, 26].

The basin-hopping method converges fast and is relatively easy to implement using common programming libraries. Since this method is heuristic by nature, we have compared its results with the tight bounds on Holevo capacity computed via a different procedure [27]. Both methods are in good agreement on all examples given in [27]. To estimate the divergence, we have run the basin-hopping computation 20 times for each channel with different initial parameters and selected the highest estimated capacity. Standard deviation after 20 runs on a set of 1000 random channels is $< \sigma_\chi > = 3.6 \cdot 10^{-4}$ and $\max \sigma_\chi = 1.1 \cdot 10^{-2}$.

To check our results in the case of unitary channels only, we have also generated random unitary channels directly by a random set of unitary matrices U_i and a random set of real coefficients c_i constrained by $\sum_i c_i^2 = 1$. Then, Kraus operators $K_i = c_i U_i$. Any such set of operators that verify $\sum_i K_i K_i^\dagger = \mathbb{1}$ defines a quantum channel. Using the Nelder-Mead method [28] with three free parameters, we have computed Holevo capacities and found that they are in full agreement with the results of the basin-hopping calculation.

V. DISCUSSION

To study the discrepancy between the capacities of the superposition and the quantum switch, note that the one-pass superposition acts as:

$$(\mathcal{C} \boxplus \mathcal{C})(\rho) = \frac{1}{2} \left(\begin{array}{c|c} \mathcal{C}(\rho) & \sum_{i,j} K_i \rho K_j^\dagger \\ \hline \sum_{i,j} K_i \rho K_j^\dagger & \mathcal{C}(\rho) \end{array}\right), \qquad (9)$$

where $\{K_i\}$ are the Kraus operators of $\mathcal{C}$ and the fixed control subsystem is omitted for brevity. To prove $\chi(\mathcal{C} \boxplus \mathcal{C}) \geq \chi(\mathcal{C})$, define a channel $\mathcal{P}$ acting on $\mathcal{L}(\mathcal{H}^c) \otimes \mathcal{L}(\mathcal{H}^t)$ with Kraus operators $\{P_m\} = \{|m\rangle\langle m|^c \otimes \mathbb{1}^t\}$. Since

$\sum_m P_m P_m^\dagger = \mathbb{1}^{ct}$, $\mathcal{P}$ is trace-preserving. It acts as:

$$\mathcal{P} \circ (\mathcal{C} \boxplus \mathcal{C})(\rho) = \mathcal{C}'(\rho) = \frac{1}{2} \left(\begin{array}{c|c} \mathcal{C}(\rho) & 0 \\ \hline 0 & \mathcal{C}(\rho) \end{array}\right). \qquad (10)$$

If the spectrum of $C(\rho)$ is $\{\eta_i\}$, then the spectrum of $C'(\rho)$ is $\{\frac{\eta_i}{2}\}$, each eigenvalue having multiplicity two. Hence:

$$\begin{aligned} H(C'(\rho)) &= -\sum_{i=1}^{8} \frac{\eta_i}{2} \log \frac{\eta_i}{2} \\ &= -\frac{1}{2} \sum_{i=1}^{8} \eta_i (\log(\eta_i) - 1) \\ &= \sum_{i=1}^{4} \eta_i \log(\eta_i) + \sum_{i=1}^{4} \eta_i \\ &= H(C(\rho)) + 1. \end{aligned} \qquad (11)$$

It now follows from (5) that:

$$\begin{aligned} \chi(\mathcal{C}') &= \max_{\{p_i,\rho_i\}} \left[H(\mathcal{C}'(\rho)) - \sum_i p_i H(\mathcal{C}'(\rho_i))\right] \\ &= \max_{\{p_i,\rho_i\}} \left[H(\mathcal{C}(\rho)) - \sum_i p_i H(\mathcal{C}(\rho_i))\right] \\ &= \chi(\mathcal{C}), \end{aligned} \qquad (12)$$

i.e. $\chi(\mathcal{P} \circ (\mathcal{C} \boxplus \mathcal{C})) = \chi(\mathcal{C})$. By the data processing inequality, a channel can only lose information between the input and the output: $\chi(\mathcal{C}) = \chi(\mathcal{P} \circ (\mathcal{C} \boxplus \mathcal{C})) \leq \chi(\mathcal{C} \boxplus \mathcal{C})$. $\square$

The quantum switch acts as:

$$\begin{aligned} &(\mathcal{C} \bowtie \mathcal{C})(\rho) = \\ &\frac{1}{2} \left(\begin{array}{c|c} \mathcal{C}(\mathcal{C}(\rho)) & \sum_{i,j} K_j K_i \rho K_j^\dagger K_i^\dagger \\ \hline \sum_{i,j} K_i K_j \rho K_i^\dagger K_j^\dagger & \mathcal{C}(\mathcal{C}(\rho)) \end{array}\right). \end{aligned} \qquad (13)$$

Unlike with quantum superposition, no simple relationship between Holevo capacities is available in this case. When the quantum switch lowers communication capacity of $\mathcal{C}$, this phenomenon has a likely origin in the loss of information in diagonal terms $\mathcal{C} \circ \mathcal{C}$. On the contrary, the advantage of the quantum switch over the superposition, when it occurs, comes from non-diagonal terms. Heuristically, the non-diagonal terms in (13) are more versatile than the non-diagonal terms in (9), explaining the behaviour of the switch in Figure 5.

Note that the self-switch $\mathcal{C} \bowtie \mathcal{C}$ is not the same channel as the superposition $(\mathcal{C} \circ \mathcal{C}) \boxplus (\mathcal{C} \circ \mathcal{C})$. If $\mathcal{C}$ has Kraus operators $\{K_i\}$, then the Kraus operators of $\mathcal{C} \circ \mathcal{C}$ are $\{K_i K_j\}$. Inserting this in (9), one obtains:

$$((\mathcal{C} \circ \mathcal{C}) \boxplus (\mathcal{C} \circ \mathcal{C}))(\rho) = \frac{1}{2} \left(\begin{array}{c|c} \mathcal{C}(\mathcal{C}(\rho)) & \mathcal{C}(\mathcal{C}(\rho)) \\ \hline \mathcal{C}(\mathcal{C}(\rho)) & \mathcal{C}(\mathcal{C}(\rho)) \end{array}\right). \qquad (14)$$

This expression is different from (13) if $\{K_i\}$ do not commute. The latter factor happens to have an even deeper
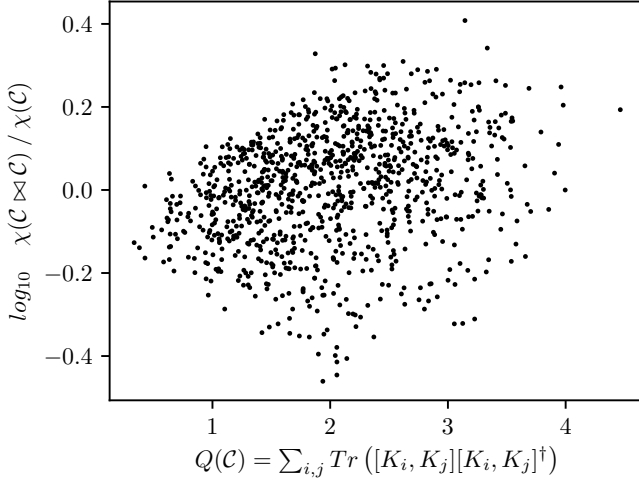
FIG. 7. Relative advantage of the quantum switch $\chi(\mathcal{C} \bowtie \mathcal{C})/\chi(\mathcal{C})$ as a function of $Q(\mathcal{C})$, degree of commutativity of the Kraus decomposition, on a set of 1000 random channels. While the quantum switch can gain or lose communication capacity depending on $\mathcal{C}$, it does so more often at higher values of $Q(\mathcal{C})$.
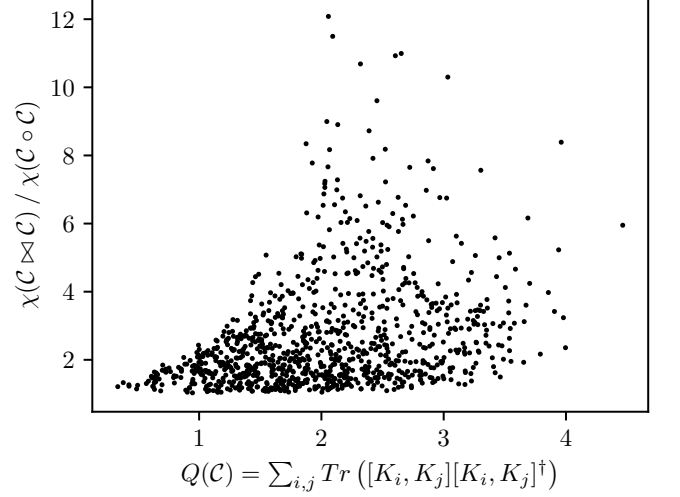

FIG. 8. Advantage $\chi(\mathcal{C} \bowtie \mathcal{C})/\chi(\mathcal{C} \circ \mathcal{C})$ as a function of $Q(\mathcal{C})$, the degree of commutativity of the Kraus decomposition. If the Kraus operators nearly commute, then the Holevo capacities are similar. If they do not commute, then the effect of the indefinite causal order gets stronger.

significance: our study of the non-diagonal terms provides evidence for the hypothesis, inspired by [13], that a greater advantage occurs when the Kraus operators do not commute. To observe this effect, define:

$$Q(\mathcal{C}) = \sum_{i,j} \mathrm{Tr}\left([K_i, K_j][K_i, K_j]^\dagger\right) \qquad (15)$$
$$= 4 - 2\mathrm{Tr}\sum_{i,j} K_i K_j K_i^\dagger K_j^\dagger.$$

$Q(\mathcal{C})$ is independent of the Kraus decomposition of $\mathcal{C}$. Indeed, note that two Kraus decompositions $\{E_i\}$ and $\{F_i\}$ describing the same channel are related by (2):

$$\sum_{i,j} E_i E_j E_i^\dagger E_j^\dagger = \sum_{i,j} \sum_{k,l,m,n} u_{ik} u_{im}^\dagger u_{jl} u_{jn}^\dagger F_k F_l F_m^\dagger F_n^\dagger$$
$$= \sum_{k,l,m,n} \delta_{km} \delta_{ln} F_k F_l F_m^\dagger F_n^\dagger$$
$$= \sum_{i,j} F_i F_j F_i^\dagger F_j^\dagger.$$

Hence $Q(\{E_i\}) = Q(\{F_i\})$. $\qquad \square$

Figures 7 and 8, generated on a set of 1000 random channels, show that the spread of $\chi(\mathcal{C} \bowtie \mathcal{C})/\chi(\mathcal{C})$ increases with Q. If $\{K_i\}$ almost commute and Q is small, then the effect of the indefinite causal order is also small: the Holevo capacity of the quantum switch is close to the Holevo capacity of the superposition. On the contrary, when Q is high, the effect of the indefinite causal order is strong: the Holevo capacity of the quantum switch is dominated by this non-commutativity and the non-diagonal terms of (13) are larger than the diagonal ones.

Our simulation shows that the same conclusion is also valid for the 3-switch [12].

VI. CONCLUSION

Both the coherent control of quantum channels and the indefinite causal order between them were shown to be resources for communicating through noisy channels. While the notion of controlling quantum channels via a quantum subsystem is a useful heuristic, its rigorous mathematical formulation shows that the action of observers involved in such control and measurement operations is hardly ever local in time and space [14, 16]. Nonetheless, as mentioned above, experimental setups have been realized for two or three subsystems combined via the quantum switch. These empirically available setups call for a better theoretical characterization of their workings.

On the one hand, it has been argued that the communication advantage found with the quantum switch "is also present" in the scenario involving one-pass superposition [15]. On the other, this advantage was hailed as a "striking ... new paradigm of Shannon theory" [13]. Our analysis via numeric simulations shows that the two resources provide the same communication gain only in very rare cases. In a more general setting, simulated by choosing either generic or unitary channels at random, the conceptual origin and the quantitative measure of the advantage differ. This is because the quantum switch exhibits an intricate interplay between two nonclassical factors, both of which are capable of providing a communication advantage: quantum superposition and

non-commutativity. What is more, the quantum switch, but not the one-pass superposition, can also induce a loss in communication capacity because it involves a composition of channels. When the switch is treated as a resource, it remains a genuine combination of different contributing properties and should not be reduced to any one of them. These results motivate further theoretical research on an important problem of quantum communication: achieving precise understanding of the interrelations between non-classical resources that provide an advantage in communication via noisy channels.

[1] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nature Communications **3**, 1092 (2012).

[2] G. Chiribella, G. M. DAriano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Physical Review A **88**, 022318 (2013).

[3] L. Hardy, Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure, Journal of Physics A: Mathematical and Theoretical **40**, 3081 (2007).

[4] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New Journal of Physics **17**, 102001 (2015).

[5] G. Chiribella and H. Kristjánsson, Quantum Shannon theory with superpositions of trajectories, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **475**, 20180903 (2019).

[6] I. Ibnouhsein and A. Grinbaum, Information-theoretic constraints on correlations with indefinite causal order, Phys. Rev. A **92**, 042124 (2015).

[7] L. M. Procopio, A. Moqanaki, M. Araújo, F. M. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, C. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nature Communications **6**, 7913 (2015).

[8] 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, **3**, e1602589 (2017), arXiv:1608.01683.

[9] 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).

[10] K. Goswami, Y. Cao, G. Paz-Silva, J. Romero, and A. White, Communicating via ignorance, (2018), arXiv:1807.07383.

[11] 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.

[12] L. M. Procopio, F. Delgado, M. Enriquez, N. Belabas, and J. A. Levenson, Communication through quantum coherent control of N channels in a multi-partite causal-order scenario, (2019), arXiv:1902.01807.

[13] D. Ebler, S. Salek, and G. Chiribella, Enhanced communication with the assistance of indefinite causal order, Phys. Rev. Lett. **120**, 120502 (2018).

[14] O. Oreshkov, Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics, arXiv e-prints (2018), 1801.07594.

[15] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, Communication through coherent control of quantum channels, (2018), arXiv:1810.09826.

[16] P. A. Guérin, G. Rubino, and C. Brukner, Communication through quantum-controlled noise, Phys. Rev. A **99**, 062317 (2019).

[17] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra and its Applications **10**, 285 (1975).

[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, NY, USA, 2011).

[19] K. Kraus, *States, Effects, and Operations. Fundamental Notions of Quantum Theory* (Springer-Verlag Berlin Heidelberg, 1983).

[20] A. S. Holevo, The capacity of the quantum channel with general signal states, IEEE Transactions on Information Theory **44**, 269 (1998).

[21] A. Holevo, *Quantum Systems, Channels, Information. A Mathematical Introduction* (De Gruyter Berlin, Boston, 2012).

[22] N. Gisin, N. Linden, S. Massar, and S. Popescu, Error filtration and entanglement purification for quantum communication, Phys. Rev. A **72**, 012338 (2005).

[23] W. Bruzda, V. Cappellini, H.-J. Sommers, and K. Życzkowski, Random quantum operations, Physics Letters A **373**, 320 (2009).

[24] D. J. Wales and J. P. K. Doye, Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms, The Journal of Physical Chemistry A **101**, 5111 (1997).

[25] R. Fletcher, *Practical Methods of Optimization; (2Nd Ed.)* (Wiley-Interscience, New York, NY, USA, 1987).

[26] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, The Journal of Chemical Physics **21**, 1087 (1953).

[27] D. Sutter, T. Sutter, P. Mohajerin Esfahani, and R. Renner, Efficient approximation of quantum channel capacities, IEEE Transactions on Information Theory **62**, 578 (2016).

[28] J. A. Nelder and R. Mead, A simplex method for function minimization, Computer Journal **7**, 308 (1965).