HAL open science

Communication through coherent control of quantum channels

Alastair Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, Cyril Branciard

To cite this version:

Alastair Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, Cyril Branciard. Communication through coherent control of quantum channels. Quantum, 2020, 4, pp.333. 10.22331/q-2020-09-24-333 . hal-02997676

HAL Id: hal-02997676

https://hal.science/hal-02997676v1

Submitted on 15 Dec 2021

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.

Communication Through Coherent Control of Quantum Channels

Alastair A. Abbott[1,2], Julian Wechs[2], Dominic Horsman[3], Mehdi Mhalla[3], and Cyril Branciard[2]

[1]Département de Physique Appliquée, Université de Genève, 1211 Genève, Switzerland

[2]Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France

[3]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble France

Submitted July 15, 2019; revised April 14, 2020; accepted September 8, 2020

arXiv:1810.09826v3 [quant-ph] 22 Sep 2020

A completely depolarising quantum channel always outputs a fully mixed state and thus cannot transmit any information. In a recent Letter [D. Ebler *et al.*, Phys. Rev. Lett. 120, 120502 (2018)], it was however shown that if a quantum state passes through two such channels in a quantum superposition of different orders—a setup known as the "quantum switch"—then information can nevertheless be transmitted through the channels. Here, we show that a similar effect can be obtained when one coherently controls between sending a target system through one of two identical depolarising channels. Whereas it is tempting to attribute this effect in the quantum switch to the indefinite causal order between the channels, causal indefiniteness plays no role in this new scenario. This raises questions about its role in the corresponding effect in the quantum switch. We study this new scenario in detail and we see that, when quantum channels are controlled coherently, information about their specific implementation is accessible in the output state of the joint control-target system. This allows two different implementations of what is usually considered to be the same channel to therefore be differentiated. More generally, we find that to completely describe the action of a coherently controlled quantum channel, one needs to specify not only a description of the channel (e.g., in terms of Kraus operators), but an additional "transformation matrix" depending on its implementation.

1 Introduction

The ability to create superpositions of quantum states opens up many advantages for communication and information processing that are inaccessible to classical mixtures of states, exemplified by their use in controlled logic gates (e.g., CNOT) in quantum computing [1]. Recently, it has been shown that a coherent quantum control system can be used to even put the causal ordering of quantum channels into superposition, thus rendering it "indefinite", in the so-called "quantum switch" [2]. Surprisingly, when certain zero-capacity channels are placed in a quantum switch the resulting switched channels still allow information to be transmitted, something impossible if their causal ordering is fixed or controlled classically [3].

Motivated by this example, we revisit here the notion of coherent control of arbitrary quantum channels—something that has in the generally been considered problematic or, at best, subtle [4–7]—by exploiting a control system to determine which channel is used to transmit a state rather than the order in which two communication channels are used. This approach of channel multiplexing has previously been used for error filtration [8] and discussed as an approach to control unknown unitaries [4]. We show here that it allows one to obtain a similar, counter-intuitive communication advantage to that obtained with the quantum switch mentioned earlier, despite the absence of any "causal indefiniteness": when each channel is maximally noisy, information can nonetheless be transmitted through the coherently multiplexed communication channels.

When controlled coherently in this way we find that—in contrast to the quantum switch [2, 3]—the action of the "global" multiplexed channel depends not only the descriptions of the individual channels as completely positive trace-preserving (CPTP) maps—i.e., their standard description as "quantum channels" [1, 9, 10]—but also on more fine-grained information about their realisations. This includes, but goes beyond, relative phase information, highlighting the subtleties involved in describing "controlled channels": indeed as we will see, without extra information on the specific channel implementation the problem is in fact ill-defined. As a result, in any framework for manipulating channels that includes their coherent control, their description must be supplemented by further information than just the corresponding CPTP map.

2 Communication through the "depolarising quantum switch"

Let us first review the quantum switch and the aforementioned communication advantage obtained exploiting it. The quantum switch is a process comprising a coherent control qubit, a d-dimensional target system, and a pair of "black box" operations that, taken individually, implement some CPTP maps—so-called "quantum channels"—$\mathcal{C}_0$ and $\mathcal{C}_1$ on their input systems [2]. If the control qubit is in the state $|0\rangle^c$, then first $\mathcal{C}_0$ then $\mathcal{C}_1$ is applied to the target system, while if it is in the state $|1\rangle^c$ then the operations are in the opposite order. Initialising the control in the state $|+\rangle^c = \frac{1}{\sqrt{2}}(|0\rangle^c + |1\rangle^c)$ therefore applies the operations in a superposition of the two orders. Since, in this case, one cannot say that either operation is definitely applied before another, the quantum switch is said to exhibit indefinite causal order [2]. The quantum switch is, more formally, a "quantum supermap" [11], and its causal indefiniteness can be formally understood through the notion of "causal non-separability" [12–15]. Several recent experiments have been conducted to implement the quantum switch and experimentally verify its causal indefiniteness [16–22].

Indefinite causal order is known to be a resource providing advantages in some tasks over any (quantum or classical) process with a definite causal order [2, 12, 23–28]. In Ref. [3], it was observed that, if the CPTP maps $\mathcal{C}_i$ are taken to be fully depolarising channels $\mathcal{N}_i$ (which map any initial target state ρ^t_{in} to the maximally mixed state $\frac{\mathbb{1}^t}{d}$), then the switch (with the initial state of the control qubit fixed to $|+\rangle^c$) implements a global channel $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$ mapping ρ^t_{in} to the joint control-target state

$$\rho^{ct}_{\text{out}} = \frac{\mathbb{1}^c}{2} \otimes \frac{\mathbb{1}^t}{d} + \frac{1}{2}\left[|0\rangle\langle 1|^c + |1\rangle\langle 0|^c\right] \otimes \frac{1}{d^2}\rho^t_{\text{in}}\,, \quad (1)$$

which is *not* $\frac{\mathbb{1}^t}{d}$ but instead retains some dependence on ρ^t_{in}. Thus, information can propagate through the "depolarising quantum switch" despite this being impossible for the channels $\mathcal{N}_0$, $\mathcal{N}_1$, $\mathcal{N}_1 \circ \mathcal{N}_0$, and $\mathcal{N}_0 \circ \mathcal{N}_1$ individually. This somewhat surprising result arising in the presence of indefinite causal order has recently been subject to experimental tests [20, 22], and generalised to setups that put more channels in a superposition of different orders [29, 30]. Note that $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$ above—and more generally, $\mathcal{S}[\mathcal{C}_0, \mathcal{C}_1]$—depend only on the CPTP maps implemented by the black-box operations, despite the fact the operations are applied only to a subspace of the joint target-control space [2, 3]; this is a consequence of the quantum switch being a quantum supermap.

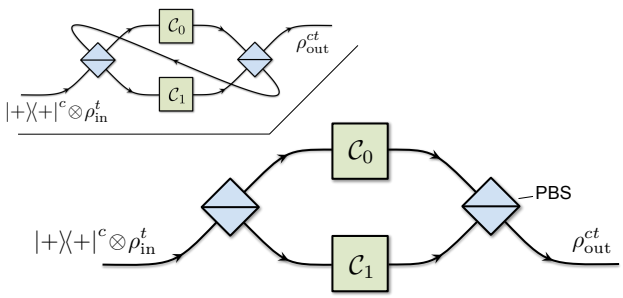

Figure 1: The upper-left inset shows a typical photonic setup for the quantum switch [2, 25], in which the control qubit is encoded in the polarisation of a photon which is routed by polarising beamsplitters (PBS), and the target system is encoded in some internal degree of freedom of the photon (as e.g. in Refs. [19, 20]). In the present work we consider only the "first half" of such a setup, as shown in the main figure. This implements a coherent control between the two boxes implementing $\mathcal{C}_0$ and $\mathcal{C}_1$ that the target system traverses. As we show, the above diagram is in fact *ill-defined* since, when controlled coherently, the maps $\mathcal{C}_i$ do not fully determine the output state ρ^{ct}_{out}; see Fig. 2.

3 Communication through coherently-controlled depolarising channels

Thus far, all attempts to implement the quantum switch have relied on interferometric setups [16–22, 31].[1] In a typical such setup, the target system is routed to the switched operations, which here correspond to communication channels, via some beamsplitters (see insert of Fig. 1). Motivated by such an approach, in this work we consider instead the state of the joint control-target system after traversing only half of such an implementation; that is, after the target system has passed, in a superposition, through the communication channels only a single time. This situation, a possible implementation of which is shown in Fig. 1, amounts to coherently controlling between applying the operations implementing $\mathcal{C}_0$ or $\mathcal{C}_1$ to the target system. By preparing the control qubit in the state $|+\rangle^c$, a "superposition" of the two operations is thus applied.[2]

We note that in general, coherent control of completely unknown quantum operations is impossible. However, in interferometer-type situations the implementations of operations are localised and thus act trivially on modes that do not pass through them; this provides additional information about the structure of the joint control-target Hilbert space that makes such

[1] While there has been some debate as to whether such experiments are genuine implementations of the quantum switch [32, 33], this issue is beyond the scope of the present paper. Here, we rather draw inspiration from such implementations to consider the case of coherently controlled channels.

[2] Our calculations below generalise easily to any other initial state of the control (not necessarily $|+\rangle^c$).

control possible [4–7]. Indeed, coherent control of unitary operations by such means has been demonstrated experimentally in many scenarios [34–38].

Let us consider, as in Refs. [3], the case where the two operations implement fully depolarising channels ($\mathcal{C}_i = \mathcal{N}_i$), and consider first the concrete case analysed in Ref. [3] and implemented experimentally in Refs. [20, 22] (for the case of a qubit target system, $d = 2$) where these are realised by randomising over a set of d^2 orthogonal unitary operators $\{U_i\}_{i=0}^{d^2-1}$. For each channel, one then indeed has $\mathcal{N}_{0/1}(\rho^t_{\text{in}}) = \frac{1}{d^2}\sum_i U_i \rho^t_{\text{in}} U_i^\dagger = \frac{\mathbb{1}^t}{d}$.

For each random choice of unitary operators (U_i, U_j), the control-target system therefore undergoes the unitary evolution $|0\rangle\langle 0|^c \otimes U_i + |1\rangle\langle 1|^c \otimes U_j$.[3] If the control qubit is initially in the state $|+\rangle^c$ and the target system is in some state $|\psi_{\text{in}}\rangle^t$, the joint system thus evolves to the state

$$|\Phi_{ij}\rangle^{ct} = \frac{1}{\sqrt{2}}\left(|0\rangle^c \otimes U_i |\psi_{\text{in}}\rangle^t + |1\rangle^c \otimes U_j |\psi_{\text{in}}\rangle^t\right). \tag{2}$$

Averaging over all choices of (U_i, U_j) one finds that the output state is

$$\begin{aligned}\rho^{ct}_{\text{out}} =& \frac{1}{d^4}\sum_{i,j} |\Phi_{ij}\rangle\langle\Phi_{ij}|^{ct} \\ =& \frac{\mathbb{1}^c}{2} \otimes \frac{\mathbb{1}^t}{d} + \frac{1}{2}\left[|0\rangle\langle 1|^c + |1\rangle\langle 0|^c\right] \otimes T\rho^t_{\text{in}}T^\dagger \end{aligned} \tag{3}$$

where $T := \frac{1}{d^2}\sum_i U_i$ and $\rho^t_{\text{in}} := |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t$. By linearity, Eq. (3) also holds for arbitrary mixed inputs ρ^t_{in}, and the setup thus gives rise to the global channel $\mathcal{M}$ mapping $\rho^t_{\text{in}} \to \rho^{ct}_{\text{out}}$.

It is immediately clear that ρ^{ct}_{out} depends in general on ρ^t_{in}, and thus some information can be transmitted through the setup. If, on the other hand, one classically controls which channel is applied to the input, no information can be transmitted. Indeed, if the initial state of the control qubit is diagonal, or if it decoheres, in the Pauli σ_z eigenbasis, then one can easily check that all dependence on ρ^t_{in} disappears in ρ^{ct}_{out}. Thus, the global channel $\mathcal{M}$ arising from coherently controlling between $\mathcal{N}_0$ and $\mathcal{N}_1$ provides a communication advantage over classical control.

This advantage mirrors that found using the quantum switch in Ref. [3], where the order in which $\mathcal{N}_0$ and $\mathcal{N}_1$ are applied is coherently controlled, thereby applying them in an indefinite causal order. In the example above, however, indefinite causal order plays no role; indeed the two channels are used in parallel (in the two arms of the interferometer), rather than one after the other. Nevertheless, the fact that coherent control—of channel or of order—plays a role in transmitting information through fully depolarising channels in both these scenarios raises questions about the role of indefinite causal order in the result of Ref. [3] using the quantum switch, or, at least, the relative role of causal indefiniteness and coherent control [39]. Similar questions, as well as the need to clarify how to compare different such advantages, have also been raised recently based on different considerations in Ref. [40], and subsequently discussed further in Ref. [41].

In Ref. [3], the authors quantified precisely how much classical information can be transmitted by a single use of the depolarising quantum switch (i.e., its Holevo information [42, 43]). In Appendix B.1 we present a lower bound for the Holevo information of the global channel $\mathcal{M}$ defined above. We find that significantly more information can be transmitted by this setup than with the full depolarising quantum switch (e.g., $\frac{1}{2}\log_2\frac{5}{4} \simeq 0.16$ bits compared to $-\frac{3}{8} - \frac{5}{8}\log_2\frac{5}{8} \simeq 0.05$ bits for a qubit target system).

It was further noted in Ref. [3] that if one traces out either the control or target system from the output of the depolarising quantum switch one obtains the completely mixed state, and thus information is transmitted solely in the correlations between the control and target states. In the present example, while it is still true that if the control is traced out the target system is left in the completely mixed state, if one traces out the target from Eq. (3) one obtains $\rho^c_{\text{out}} = \frac{1}{2}(\mathbb{1}^c + \text{Tr}[T\rho^t_{\text{in}}T^\dagger]\sigma^c_x)$, which still depends on ρ^t_{in}. Nevertheless, the control system itself does not contain (at any stage in the interferometer) all of the information about the input target state that gets transmitted to ρ^{ct}_{out} (it only contains $\text{Tr}[T\rho^t_{\text{in}}T^\dagger]$, while ρ^{ct}_{out} contains $T\rho^t_{\text{in}}T^\dagger$). Note that in the setup of Fig. 1, just as in the depolarising quantum switch, the subspace on which each channel acts is, technically, a nontrivial d-dimensional "sector" of he $2d$-dimensional joint control-target Hilbert space[4] [44], and so the channels cannot be strictly said to act trivially on the control system [32]. Indeed, in both cases the global channel acts to delocalise some of the information about the target into the correlations between the two systems. This is conceptually similar to the effect of quantum phase kickback associated with controlled unitary operations [1, 45], and its role in such communication advantages is further explored in Refs. [40, 41, 44].

4 Dependence on channel implementation

The approach employed above of randomising over unitary channels is not, however, the only way to implement a fully depolarising channel. Recall that in

[3] Note that U_i and U_j (and similarly, the Kraus operators K_i and L_j considered later on) must be written with respect to a common reference phase. We also assume that the arms of the circuit do not introduce any additional relative phases).

[4] That is, the channel $\mathcal{C}_i$ acts on the sector corresponding to the control system being in the state $|i\rangle\langle i|^c$.

general, a quantum channel $\mathcal{C}$ is defined as a CPTP map, and can be described in terms of a (non-unique) set of Kraus operators $\{K_i\}_i$ satisfying $\sum_i K_i^\dagger K_i = \mathbb{1}$, such that the output of the channel is given by $\mathcal{C}(\rho) = \sum_i K_i \rho K_i^\dagger$ for every density matrix ρ [1, 9, 10]. Note in particular that $\mathcal{C}$ may be applied to a subsystem in a subspace of some larger Hilbert space, as is the case both in the quantum switch and the scenario of Fig. 1. There, however, if the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ are not unitary—or not described, as previously considered, as a randomisation over unitary channels—it is *a priori* unclear how to determine the global channel mapping $\rho^t_{\text{in}} \to \rho^{ct}_{\text{out}}$ from the Kraus operators of $\mathcal{C}_0$ and $\mathcal{C}_1$.

One possible approach to doing so is to "purify" the channels via (independent) Stinespring dilations [46]. Any channel $\mathcal{C}$ with Kraus operators $\{K_i\}_i$ can indeed be extended to a unitary operation by introducing an environment in an initial state $|\varepsilon\rangle^e$ and considering the operation that acts on the system under consideration (in our case, the target) and the environment as $|\psi_{\text{in}}\rangle^t \otimes |\varepsilon\rangle^e \to \sum_i K_i |\psi_{\text{in}}\rangle^t \otimes |i\rangle^e := |\Phi_{\text{out}}\rangle^{te}$, where the ket vectors $|i\rangle^e$ are (normalised) orthogonal states of the environment. After tracing out the environment, we recover $\text{Tr}_e |\Phi_{\text{out}}\rangle\langle\Phi_{\text{out}}|^{te} = \sum_i K_i |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t K_i^\dagger = \mathcal{C}(|\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t)$, as required.

In the setup of Fig. 1 where the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ have Kraus operators $\{K_i\}_i$ and $\{L_j\}_j$, respectively, one may therefore purify the channels by introducing two, initially uncorrelated, environments with initial states $|\varepsilon_0\rangle^{e_0}$ and $|\varepsilon_1\rangle^{e_1}$. Note that the control qubit must then be seen as controlling the action of the purified unitary extensions of the channels not only on the target system, but also on the corresponding environments. This is nevertheless sensible in the interferometric picture of Fig. 1 where the channels may be seen as black boxes with "internal" environments, that a photon traverses (in a superposition of "here" and "there").

Under these controlled, purified channels, the combined control-target-environments state evolves unitarily as

$$\begin{aligned} &|+\rangle^c \otimes |\psi_{\text{in}}\rangle^t \otimes |\varepsilon_0\rangle^{e_0} \otimes |\varepsilon_1\rangle^{e_1} \\ &\to \frac{1}{\sqrt{2}} |0\rangle^c \otimes \sum_i K_i |\psi_{\text{in}}\rangle^t \otimes |i\rangle^{e_0} \otimes |\varepsilon_1\rangle^{e_1} \\ &\quad + \frac{1}{\sqrt{2}} |1\rangle^c \otimes \sum_j L_j |\psi_{\text{in}}\rangle^t \otimes |\varepsilon_0\rangle^{e_0} \otimes |j\rangle^{e_1} . \end{aligned} \tag{4}$$

After tracing out the environments, the resulting joint control-target state ρ^{ct}_{out} is found to be

$$\begin{aligned} \rho^{ct}_{\text{out}} =& \frac{1}{2}\left[|0\rangle\langle 0|^c \otimes \mathcal{C}_0(\rho^t_{\text{in}}) + |1\rangle\langle 1|^c \otimes \mathcal{C}_1(\rho^t_{\text{in}})\right] \\ &+ \frac{1}{2}\left[|0\rangle\langle 1|^c \otimes T_0 \rho^t_{\text{in}} T_1^\dagger + |1\rangle\langle 0|^c \otimes T_1 \rho^t_{\text{in}} T_0^\dagger\right] \end{aligned} \tag{5}$$

with $T_0 := \sum_i \langle\varepsilon_0|i\rangle K_i$ and $T_1 := \sum_j \langle\varepsilon_1|j\rangle L_j$.

The output state (3), obtained when $\mathcal{C}_0$ and $\mathcal{C}_1$ are depolarising channels implemented as a classical randomisation over d^2 orthogonal unitary operators U_i, is recovered by taking $K_i = \frac{1}{d} U_i$, $L_j = \frac{1}{d} U_j$, and the initial states of the environment to be $|\varepsilon_0\rangle^{e_0} = \sum_{i=0}^{d^2-1} \frac{1}{d} |i\rangle^{e_0}$, $|\varepsilon_1\rangle^{e_1} = \sum_{j=0}^{d^2-1} \frac{1}{d} |j\rangle^{e_1}$. Note, however, that a different choice of orthogonal unitary operations (even due to the addition of a relative phase between them, so that, taken individually, they would still implement the same local unitary channels) would have led to a different output state in Eq. (5). If we had instead taken the environments to initially be in the states $|0\rangle^{e_0}$ and $|0\rangle^{e_1}$ and chosen a set of orthogonal unitary operators such that $K_0 = L_0 = \frac{1}{d}\mathbb{1}$, we would have obtained Eq. (3) with $T = \frac{1}{d}\mathbb{1}$—which, incidentally, coincides with the state of Eq. (1) obtained in Ref. [3] as the output of the depolarising quantum switch. We nevertheless emphasise that Eq. (5) gives the output control-target state not only when the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ are obtained by classical randomisation over unitary channels, but for any description of the channels in terms of Kraus operators.

The crucial observation here is that ρ^{ct}_{out} depends on the implementation of the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ [47, 48]. The interferometric circuit in Fig. 1 is therefore not fully defined by the channels $\mathcal{C}_0$ and $\mathcal{C}_1$, or the Kraus operators chosen to represent them. This may appear surprising given that, in the usual paradigm, quantum channels are understood to be fully characterised by their (non-unique) Kraus representation [1, 9, 10]. However, one should note that such a description of a channel is unchanged under addition of any global phase. On the other hand, any such "global" phase applied by one of the channels in Fig. 1 is only applied to the corresponding arm of the interferometer and therefore, in the overall controlled circuit, becomes a "relative" phase with physical significance. In the case where the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ are unitary, the fact that Fig. 1 is only defined up to such a phase on the unitaries is well known [36, 45].

What we see here, however, is that the output of the interferometric circuit depends not only on any relative phases between (the Kraus operators of) the two channels, but also on a more detailed description of the implementation of the channels. More precisely, one requires some additional information encoded in the matrices T_0, T_1 introduced in Eq. (5) in order to fully specify the global channel $\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1] : \rho^t_{\text{in}} \to \rho^{ct}_{\text{out}}$ induced by the circuit; see Fig. 2. We call these the "transformation matrices" of the channel implementations. In the description above in terms of a Stinespring dilation, these depend not only on the set of Kraus operators used to decompose the channel, but also on how these are combined (with coefficients that depend on the environment states)

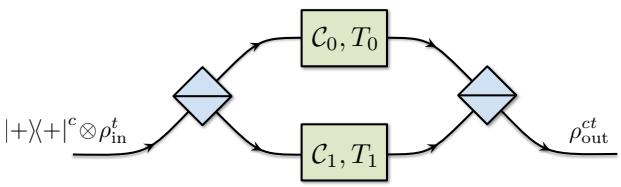

Figure 2: A corrected version of Fig. 1 in which the description of the two operations inside the interferometer, implementing the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ on their respective subspaces, have been supplemented by the transformation matrices T_0 and T_1 needed to fully specify the output state ρ^{ct}_{out}.

to define $T_0 := \sum_i \langle \varepsilon_0 | i \rangle K_i$ and $T_1 := \sum_j \langle \varepsilon_1 | j \rangle L_j$.[5] Let us emphasise that as any channel can be purified in a form that is equivalent to a Stinespring dilation [1, 9, 10], the description above is fully general.

In Appendix A (see Eq. (A9)), we characterise completely the transformation matrices T obtainable from some realisation of any given channel $\mathcal{C}$, by deriving a general constraint expressed in terms of the Choi representations [49] of $\mathcal{C}$ and T. For a d-dimensional fully depolarising channel, for instance, this constraint simplifies to $\mathrm{Tr}[T^\dagger T] \leq \frac{1}{d}$. Under this constraint, applied to both T_0 and T_1, Eq. (5) characterises all possible output states that one can obtain from the setup of Fig. 2, for any implementation of the channels $\mathcal{C}_0, \mathcal{C}_1 = \mathcal{N}$.[6]

Finally, recalling that the quantum switch is a "quantum supermap" [11] we note that the transformation $\mathcal{S}[\mathcal{C}_0, \mathcal{C}_1]$ it induces has no such dependence on the implementation of $\mathcal{C}_0$ and $\mathcal{C}_1$, something which is more generally true in any setup in which each channel is applied once and only once to the target system irrespective of the state of the control (see Appendix C). The coherent control of quantum channels is thus not a quantum supermap. We note, however, that Ref. [44] has recently introduced the notion of "vacuum-extended channels" as an alternative approach to capturing the implementation-dependence relevant to the coherent control of quantum channels. There, coherent control can be expressed as a supermap of the *vacuum-extended* channels corresponding to a particular implementation, allowing one to formally prove the absence of indefinite causal order (i.e., in that case, the causal separability) of the coherent control of such channels.

[5]Note that while T_0 and T_1 are both independently defined with respect to a common phase reference (supposed to be shared by the two "boxes" containing the channels $\mathcal{C}_0, \mathcal{C}_1$ and their environments), only their relative phase is in fact needed to obtain the output state ρ^{ct}_{out}.

[6]Note that this implies that for certain implementations of depolarising channels the global channel $\mathcal{M}[\mathcal{N}_0, T_0, \mathcal{N}_1, T_1]$ cannot transmit any information; indeed, this is the case if $T_0 = 0$ or $T_1 = 0$.

5 Distinguishing different implementations of coherently-controlled channels

The dependence of the output of the circuit of Fig. 2 on the implementation of the channels means that it is also possible to differentiate between two distinct implementations of the same quantum channel with different transformation matrices.

To see this, consider the case where the channel $\mathcal{C}_0$ has a single, fixed implementation with a transformation matrix T_0, while the channel $\mathcal{C}_1$ can have two different possible implementations, with $T_1 \neq T_1'$. The global channels $\mathcal{M}_{T_1} := \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1]$ and $\mathcal{M}_{T_1'} := \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1']$ thus differ in general. If T_1 and T_1' are equally probable, then the maximal probability of successfully distinguishing the two channels—and thereby the two implementations of $\mathcal{C}_1$—is $\frac{1}{2}(1 + \mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}))$, where $\mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}) := \frac{1}{2}\|\mathcal{M}_{T_1} - \mathcal{M}_{T_1'}\|_\diamond$ is the diamond-norm distance between the two global channels [50]. In Appendix D we show that

$$\mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}) \leq \frac{1}{2}\left\|T_1 - T_1'\right\|_2 \tag{6}$$

(where $\|\cdot\|_2$ is the spectral norm), and that this upper bound can be reached with $\mathcal{C}_0 = \mathcal{I}$, $T_0 = \mathbb{1}$, in which case it is obtained by taking the input state $\rho^t_{\text{in}} = |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|$ maximising $\langle\psi_{\text{in}}| (T_1 - T_1')^\dagger (T_1 - T_1') |\psi_{\text{in}}\rangle$. One then discriminates the channels by performing optimal state discrimination between the corresponding output states ρ^{ct}_{out} and $\rho^{ct\,\prime}_{\text{out}}$ of the two global channels.

It is indeed well-known, for instance, that the interferometric setup of Fig. 2 allows one to perfectly discriminate whether the lower arm applies the operation $|\psi\rangle^t \to |\psi\rangle^t$ or $|\psi\rangle^t \to -|\psi\rangle^t$ (the unitaries $\pm\mathbb{1}$), even though these both correspond to the identity channel $\mathcal{C}_1 = \mathcal{I}$ (but with $T_1, T_1' = \pm\mathbb{1}$) on the relevant subspace. As another, perhaps more interesting, example consider the case where $\mathcal{C}_1 = \mathcal{N}$ is the fully depolarising channel, with the two possible transformation matrices $T_1^{(\prime)} = \pm\frac{1}{\sqrt{d}}|0\rangle\langle 0|$. We have $\frac{1}{2}\left\|T_1 - T_1'\right\|_2 = \frac{1}{\sqrt{d}}$, so that these two implementations of the depolarising channel can be distinguished (for $\mathcal{C}_0 = \mathcal{I}$, $T_0 = \mathbb{1}$, $\rho^t_{\text{in}} = |0\rangle\langle 0|^t$) with probability $\frac{1}{2}(1 + \frac{1}{\sqrt{d}})$ ($\simeq 0.85$ for $d = 2$).

6 Discussion

Coherent control of quantum channels was previously shown to be a resource for communicating through noisy channels in the technique of "error filtration" [8]. Our analysis, following that of Ref. [3], shows how coherent control of channels provides more general communication advantages, increasing the capacity of transmission in the absence of postselection and even in the extreme case of completely depolarising

channels. This information gain could be useful in more general error correction and mitigation scenarios, both for quantum communication and computation.

While we focused on depolarising channels to illustrate the ability for two coherently controlled zero-capacity channels to transmit information, this is not the only case where one should intuitively expect no communication to be possible. Indeed, any constant channel has zero capacity [42, 43] and similar conclusions can be drawn for any such channel. Furthermore, while this situation allows for the communication of *classical* information, we note that the authors of Ref. [3] also recently investigated the transmission of *quantum* information through a quantum switch that puts two dephasing channels in a superposition of orders [51]. In fact the advantage found there with the quantum switch is also present in the analogous scenario of Fig. 2; see Appendix B.2. Nevertheless, a related recent result has also shown that, for certain choices of noisy channels, the quantum switch can enable *perfect* quantum communication, but that this impossible with our setup [52]. Understanding when the quantum switch or the coherent control of channels is a more powerful communication resource [39], as well as the role of indefinite causal order in communication advantages provided by the quantum switch [41], are crucial questions that remain the be fully answered.

Our analysis illuminated the fact that the global transformation implemented by the circuit in Fig. 2 depends on the implementation of whatever channels are used. This stands in contrast to the usual paradigm of quantum channels, where a channel is defined as a CPTP map, and where all descriptions in terms of Kraus operators, or all purifications of a quantum channel, are equivalent [1, 9, 10]. Although such a description suffices if a channel is only ever used in isolation, by exploiting quantum control (something possible when the channel is supplied as a "black-box" or a usable communication channel [4]) it is in fact possible to extract information about how a channel is implemented, opening up the possibility to use coherent control as a tool for, e.g., error correction [53], quantum channel security [54] and characterisation [48].

Our results thus show that the notions of coherently controlling quantum channels—and, by extension, their actions when composed in circuits—is, by itself, ill-defined. Nevertheless, the setups in Figs. 1–2 that we have considered in this work are perfectly realisable experimentally; indeed, they are in fact less demanding than the approaches already used to implement the full quantum switch [16–22]. Our observations here add to the call (e.g., in Ref. [4] for the control of unknown unitaries) for a generalisation of the standard paradigm of quantum circuits to describe experimentally conceivable situations, that would include the possibility for operations to be quantum-controlled (a general quantum "if statement"), or more generally to be applied on subspaces only. In the situation we considered, we saw that (generalised) quantum channels could be defined not only by the CPTP maps they induce, but also required one to specify the "transformation matrices" T introduced above. We expect that this approach can be used (possibly as a complement to that recently proposed in Ref. [44]) for more general situations than the one investigated here [55], and leave its possible generalisation for future work.

Acknowledgements We thank Časlav Brukner, Giulio Chiribella, Philippe Allard Guérin, and Hlér Kristjánsson for comments on this manuscript. We acknowledge financial support from the Swiss National Science Foundation (NCCR SwissMAP and Starting Grant DIAQ) and the *"Investissements d'avenir"* (ANR-15-IDEX-02) program of the French National Research Agency.

References

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, NY, USA, 2011).

[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] 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].

[4] M. Araújo, A. Feix, F. Costa, and Č. Brukner, Quantum circuits cannot control unknown operations, New J. Phys. **16**, 093026 (2014), arXiv:1309.7976 [quant-ph].

[5] N. Friis, V. Dunjko, W. Dür, and H. J. Briegel, Implementing quantum control for unkown subroutines, Phys. Rev. A **89**, 030303(R) (2014), arXiv:1401.8128 [quant-ph].

[6] T. M. Rambo, J. B. Altepeter, P. Kumar, and G. M. D'Ariano, Functional quantum computing: An optical approach, Phys. Rev. A **93**, 052321 (2016), arXiv:1211.1257 [quant-ph].

[7] J. Thompson, K. Modi, V. Vedral, and M. Gu, Quantum plug n' play: modular computation in the quantum regime, New J. Phys. **20**, 013004 (2018), arXiv:1310.2927 [quant-ph].

[8] N. Gisin, N. Linden, S. Massar, and S. Popescu, Error filtration and entanglement purification for quantum communication, Phys. Rev. A **72**, 012338 (2005), arXiv:quant-ph/0407021.

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

[10] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013) arXiv:1106.1445 [quant-ph].

[11] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, EPL **83**, 30004 (2008), arXiv:0804.0180 [quant-ph].

[12] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].

[13] 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].

[14] O. Oreshkov and C. Giarmatzi, Causal and causally separable processes, New J. Phys. **18**, 093020 (2016), arXiv:1506.05449 [quant-ph].

[15] 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].

[16] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nat. Commun. **6**, 7913 (2015), arXiv:1412.4006 [quant-ph].

[17] 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].

[18] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, Experimental entanglement of temporal orders (2017), arXiv:1712.06884 [quant-ph].

[19] 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].

[20] K. Goswami, J. Romero, and A. G. White, Communicating via ignorance: Increasing communication capacity via superposition of order, Phys. Rev. Research **2**, 033292 (2018), arXiv:1807.07383 [quant-ph].

[21] 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].

[22] 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].

[23] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301 (2012), arXiv:1109.5154 [quant-ph].

[24] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti, Quantum computation with programmable connections between gates, Phys. Lett. A **376**, 2940 (2012), arXiv:1109.5987 [quant-ph].

[25] 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].

[26] S. Facchini and S. Perdrix, Quantum circuits for the unitary permutation problem, in *TAMC 2015: Theory and Applications of Models of Computation*, edited by R. Jain, S. Jain, and F. Stephan (Springer International Publishing, Cham, 2015) pp. 324–331, arXiv:1405.5205 [quant-ph].

[27] 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].

[28] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, Exponential communication complexity advantage from quantum superposition of the direction of communication, Phys. Rev. Lett. **117**, 100502 (2016), arXiv:1605.07372 [quant-ph].

[29] L. M. Procopio, F. Delgado, M. Enriquez, N. Belabas, and J. A. Levenson, Communication enhancement through quantum coherent control of N channels in an indefinite causal-order scenario, Entropy **21**, 1012 (2019), arXiv:1902.01807 [quant-ph].

[30] L. M. Procopio, F. Delgado, M. Enriquez, N. Belabas, and J. A. Levenson, Sending classical information via three noisy channels in superposition of causal orders, Phys. Rev. A **101**, 012346 (2020), arXiv:1910.11137 [quant-ph].

[31] M. M. Taddei, J. C. ne, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, Experimental computational advantage from superposition of multiple temporal orders of quantum gates (2020), arXiv:2002.07817 [quant-ph].

[32] O. Oreshkov, Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics, Quantum **3**, 206 (2019), arXiv:1801.07594 [quant-ph].

[33] N. Paunkovic and M. Vojinovic, Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders, Quantum **4**, 275 (2020), arXiv:1905.09682 [quant-ph].

[34] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White, Simplifying quantum logic using higher-dimensional Hilbert spaces, Nat. Phys. **5**, 134 (2009), arXiv:0804.0272 [quant-ph].

[35] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien, Adding control to arbitrary unknown quantum operations, Nat. Commun. **2**, 413 (2011), arXiv:1006.2670 [quant-ph].

[36] X.-Q. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien, Calculating unknown eigenvalues with a quantum algorithm, Nat. Photonics **7**, 223 (2013), arXiv:1110.4276 [quant-ph].

[37] N. Friis, A. A. Melnikov, G. Kirchmair, and H. J. Briegel, Coherent controlization using superconducting qubits, Sci. Rep. **5**, 18036 (2015), arXiv:1508.00447 [quant-ph].

[38] V. Dunjko, N. Friis, and H. J. Briegel, Quantum-enhanced deliberation of learning agents using trapped ions, New J. Phys. **17**, 023006 (2015), arXiv:1407.2830 [quant-ph].

[39] N. Loizeau and A. Grinbaum, Channel capacity enhancement with indefinite causal order, Phys. Rev. A **101**, 012340 (2020), arXiv:1906.08505 [quant-ph].

[40] P. A. Guérin, G. Rubino, and Č. Brukner, Communication through quantum-controlled noise, Phys. Rev. A **99**, 062317 (2019), arXiv:1812.06848 [quant-ph].

[41] H. Kristjánsson, G. Chiribella, S. Salek, D. Ebler, and M. Wilson, Resource theories of communication with quantum superpositions of processes, New J. Phys. **22**, 073014 (2020), arXiv:1910.08197 [quant-ph].

[42] B. Schumacher and M. D. Westmoreland, Sending classical information via noisy quantum channels, Phys. Rev. A **56**, 131 (1997).

[43] A. S. Holevo, The capacity of the quantum channel with general signal states, IEEE Trans. Inf. Theory **44**, 269 (1998), arXiv:quant-ph/9611023.

[44] G. Chiribella and H. Kristjánsson, Quantum Shannon theory with superpositions of trajectories, Proc. R. Soc. A **475**, 20180903 (2019), arXiv:1812.05292 [quant-ph].

[45] A. Bisio, M. Dall'Arno, and P. Perinotti, Quantum conditional operations, Phys. Rev. A **94**, 022340 (2016), arXiv:1509.01062 [quant-ph].

[46] W. F. Stinespring, Positive functions on C^*-algebras, Proc. Amer. Math. Soc. **6**, 211 (1955).

[47] J. Åberg, Subspace preservation, subspace locality, and gluing of completely positive maps, Ann. Phys. **313**, 326 (2004), arXiv:quant-ph/0302182.

[48] D. K. L. Oi, Interference of quantum channels, Phys. Rev. Lett. **91**, 067902 (2003), arXiv:quant-ph/0303178.

[49] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).

[50] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).

[51] S. Salek, D. Ebler, and G. Chiribella, Quantum communication in a superposition of causal orders (2018), arXiv:1809.06655 [quant-ph].

[52] 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 channel (2018), arXiv:1810.10457 [quant-ph].

[53] S. J. Devitt, W. J. Munro, and K. Nemoto, Quantum error correction for beginners, Rep. Prog. Phys. **76**, 076001 (2013), arXiv:0905.2794 [quant-ph].

[54] A. Ambainis, M. Mosca, A. Tapp, and R. De Wolf, Private quantum channels, in *Proc. 41st Annual Symposium on Foundations of Computer Science* (IEEE, 2000) pp. 547–553.

[55] Q. Dong, S. Nakayama, A. Soeda, and M. Murao, Controlled quantum operations and combs, and their applications to universal controllization of divisible unitary operations (2019), arXiv:1911.01645 [quant-ph].

[56] J. C. A. Barata and M. S. Hussein, The Moore–Penrose pseudoinverse: A tutorial review of the theory, Braz. J. Phys. **42**, 146 (2012), arXiv:1110.6882 [math-ph].

[57] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A **55**, 1613 (1997), arXiv:quant-ph/9604015.

[58] P. W. Shor, The quantum channel capacity and coherent information, in *Lecture notes, MSRI Workshop on Quantum Computation* (2002).

[59] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Trans. Inf. Theory **51**, 44 (2005), arXiv:quant-ph/0304127.

[60] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, in preparation.

A Characterisation of the possible transformation matrices T of a channel $\mathcal{C}$

Here we characterise completely the transformation matrices T that can be obtained by some implementation of a given channel $\mathcal{C}$, before presenting some examples for specific channels of interest. To this end, we first recall some details about the Choi isomorphism which will allow us to concisely state and prove our characterisation.

A.1 Choi isomorphism

For any given operator $T : \mathcal{H}_I \to \mathcal{H}_O$ from some input Hilbert space $\mathcal{H}_I$ to some output Hilbert space $\mathcal{H}_O$ (which, for simplicity, we both take to be finite-dimensional), one can define its *Choi vector* representation [49] as

$$|T\rangle\!\rangle \coloneqq \mathbb{1} \otimes T\,|\mathbb{1}\rangle\!\rangle = \sum_m |m\rangle \otimes T\,|m\rangle \;\in \mathcal{H}_I \otimes \mathcal{H}_O, \tag{A1}$$

where $\{|m\rangle\}_m$ is a fixed orthonormal basis of $\mathcal{H}_I$ and $|\mathbb{1}\rangle\!\rangle \coloneqq \sum_m |m\rangle \otimes |m\rangle$. Reciprocally, it is easy to see that given a Choi vector $|T\rangle\!\rangle \in \mathcal{H}_I \otimes \mathcal{H}_O$, one can recover its corresponding operator $T : \mathcal{H}_I \to \mathcal{H}_O$ as

$$T = \sum_{m,n} \langle m, n\,|T\rangle\!\rangle\, |n\rangle\langle m|, \tag{A2}$$

where $\{|n\rangle\}_n$ is now a fixed orthonormal basis of $\mathcal{H}_O$, and where $\langle m, n| = \langle m| \otimes \langle n|$.

In a similar way, one can define the *Choi matrix* representation of any given channel $\mathcal{C}$ from $\mathcal{L}(\mathcal{H}_I)$ to $\mathcal{L}(\mathcal{H}_O)$ (with $\mathcal{L}(\mathcal{H})$ denoting the space of linear operators over the Hilbert space $\mathcal{H}$) as

$$C \coloneqq \mathcal{I} \otimes \mathcal{C}(|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|) = \sum_{m,m'} |m\rangle\langle m'| \otimes \mathcal{C}(|m\rangle\langle m'|) \in \mathcal{L}(\mathcal{H}_I \otimes \mathcal{H}_O), \tag{A3}$$

where $\mathcal{I}$ is the identity channel. The channel $\mathcal{C}$ is recovered from its Choi matrix as follows:

$$\mathcal{C}(\rho) = \mathrm{Tr}_{\mathcal{H}_I}[C \cdot (\rho^{\mathrm{T}} \otimes \mathbb{1}^{\mathcal{H}_O})] \quad \forall\, \rho, \tag{A4}$$

where $\mathrm{Tr}_{\mathcal{H}}$ generically denotes the partial trace over a Hilbert space $\mathcal{H}$, T denotes transposition in the chosen basis $\{|m\rangle\}_m$, and where for clarity the superscript in $\mathbb{1}^{\mathcal{H}_O}$ indicates the Hilbert space on which the identity operator acts. The fact that $\mathcal{C}$ is, by definition, a completely positive map ensures that its Choi matrix C is positive semidefinite; the fact that $\mathcal{C}$ is trace-preserving implies that $\mathrm{Tr}_{\mathcal{H}_O} C = \mathbb{1}^{\mathcal{H}_I}$.

One can easily see that if the channel $\mathcal{C}$ is characterised by a set of Kraus operators $\{K_i\}_i$ (i.e., such that $\mathcal{C}(\rho) = \sum_i K_i \rho K_i^\dagger \;\forall \rho$), then their Choi representations satisfy

$$C = \sum_i |K_i\rangle\!\rangle\langle\!\langle K_i|\,. \tag{A5}$$

Written in this form, it is indeed clear that C is Hermitian positive semidefinite; the constraint that $\mathrm{Tr}_{\mathcal{H}_O} C = \mathbb{1}^{\mathcal{H}_I}$ is equivalent to $\sum_i K_i^\dagger K_i = \mathbb{1}$. We further note that all the $|K_i\rangle\!\rangle$'s are necessarily in the range of C: $|K_i\rangle\!\rangle \in \mathrm{range}(C)\;\forall\, i$.[7]

A.2 General constraints on the transformation matrices

Consider a channel $\mathcal{C}$, with its Choi matrix C, and let us denote by C^+ the Moore-Penrose pseudoinverse of C [56]. As C is positive semidefinite, its pseudoinverse is also positive semidefinite and can be obtained as follows: diagonalising C in the form[8] $C = \sum_k |C_k\rangle\!\rangle\langle\!\langle C_k|$, with nonzero orthogonal eigenvectors $|C_k\rangle\!\rangle$, one simply has $C^+ = \sum_k \frac{|C_k\rangle\!\rangle\langle\!\langle C_k|}{\langle\!\langle C_k|C_k\rangle\!\rangle^2}$.

In the case where $T = \sum_i \langle\varepsilon_0|i\rangle K_i$ is the transformation matrix for a given implementation of the channel $\mathcal{C}$,

[7] This can easily be seen, e.g., by introducing the projector $\Pi_{\mathcal{C}}^\perp$ onto the orthogonal complement of C, and by noting that $\mathrm{Tr}(\Pi_{\mathcal{C}}^\perp C) = 0 = \sum_i \mathrm{Tr}(\Pi_{\mathcal{C}}^\perp\, |K_i\rangle\!\rangle\langle\!\langle K_i|)$ implies (as each $\mathrm{Tr}(\Pi_{\mathcal{C}}^\perp\, |K_i\rangle\!\rangle\langle\!\langle K_i|) \geq 0$) that $\Pi_{\mathcal{C}}^\perp\, |K_i\rangle\!\rangle = 0$ for all i.

[8] Note that this diagonalisation of C is of the form of Eq. (A5), so that the operators $C_k = \sum_{m,n} \langle m, n\,|C_k\rangle\!\rangle\, |n\rangle\langle m|$ thus obtained define valid ("canonical") Kraus operators for the channel $\mathcal{C}$.

we have that its Choi vector $|T\rangle\!\rangle = \sum_i \langle\varepsilon_0|i\rangle\, |K_i\rangle\!\rangle \in \mathrm{range}(C)$ satisfies

$$\begin{aligned}
\langle\!\langle T|\, C^+\, |T\rangle\!\rangle^2 &= \Big|\sum_i \langle\varepsilon_0|i\rangle\, \langle\!\langle T|\, C^+\, |K_i\rangle\!\rangle\Big|^2 \\
&\leq \Big(\sum_i \big|\langle\varepsilon_0|i\rangle\big|^2\Big)\Big(\sum_i \big|\,\langle\!\langle T|\, C^+\, |K_i\rangle\!\rangle\,\big|^2\Big) \\
&\leq \sum_i \langle\!\langle T|\, C^+\, |K_i\rangle\!\rangle\langle\!\langle K_i|\, C^+\, |T\rangle\!\rangle \\
&= \langle\!\langle T|\, C^+ C C^+\, |T\rangle\!\rangle = \langle\!\langle T|\, C^+\, |T\rangle\!\rangle\,,
\end{aligned} \tag{A6}$$

where in the second line we used the Cauchy-Schwartz inequality, in the third line we used the fact (due to the normalisation of $|\varepsilon_0\rangle$) that $\sum_i |\langle\varepsilon_0|i\rangle|^2 \leq 1$, and in the fourth line we made use of Eq. (A5) and of the fact that the Moore-Penrose pseudoinverse satisfies $C^+CC^+ = C^+$. From Eq. (A6) it then follows that

$$\langle\!\langle T|\, C^+\, |T\rangle\!\rangle \leq 1. \tag{A7}$$

Conversely, suppose that an operator T satisfies $|T\rangle\!\rangle \in \mathrm{range}(C)$ and $\langle\!\langle T|\, C^+\, |T\rangle\!\rangle \leq 1$. We will see that such a T is the transformation matrix obtained from a particular implementation of the channel $\mathcal{C}$. Consider indeed the diagonalisation $C = \sum_k |C_k\rangle\!\rangle\langle\!\langle C_k|$ introduced already, and define the coefficients $\varepsilon_k := \frac{\langle\!\langle C_k|T\rangle\!\rangle}{\langle\!\langle C_k|C_k\rangle\!\rangle}$. By assumption one has $\sum_k |\varepsilon_k|^2 = \langle\!\langle T|\, \sum_k \frac{|C_k\rangle\!\rangle\langle\!\langle C_k|}{\langle\!\langle C_k|C_k\rangle\!\rangle^2}\, |T\rangle\!\rangle = \langle\!\langle T|\, C^+\, |T\rangle\!\rangle \leq 1$, so that the ε_k's define valid (subnormalised) coefficients, allowing us to define the initial state $|\varepsilon_0\rangle$ of the environment such that $\langle\varepsilon_0|k\rangle = \varepsilon_k$. One then finds

$$\sum_k \langle\varepsilon_0|k\rangle\, |C_k\rangle\!\rangle = \sum_k \frac{|C_k\rangle\!\rangle\langle\!\langle C_k|}{\langle\!\langle C_k|C_k\rangle\!\rangle}\, |T\rangle\!\rangle = CC^+\, |T\rangle\!\rangle = |T\rangle\!\rangle \tag{A8}$$

(where we used the assumption that $|T\rangle\!\rangle \in \mathrm{range}(C)$ and the fact that CC^+ is the orthogonal projector onto $\mathrm{range}(C)$). Equivalently, $\sum_k \langle\varepsilon_0|k\rangle C_k = T$, so we see that the transformation matrix for the particular implementation of the channel $\mathcal{C}$ obtained from the Kraus operators $\{C_i\}_i$ and the initial state $|\varepsilon_0\rangle$ of the environment specified above, and using the Stinespring dilation technique as in the main text is indeed T, as desired.

From these observations, one can thus characterise the set $\mathcal{T}_\mathcal{C}$ of all possible transformation matrices T of a given channel $\mathcal{C}$ as

$$\mathcal{T}_\mathcal{C} = \big\{T : |T\rangle\!\rangle \in \mathrm{range}(C) \text{ and } \langle\!\langle T|\, C^+\, |T\rangle\!\rangle \leq 1\big\}. \tag{A9}$$

A.3 Examples

The Choi matrix of an identity channel $\mathcal{I}$ is $I = |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|$ (in any dimension); its range is the span of $|\mathbb{1}\rangle\!\rangle$ only, and its Moore-Penrose pseudoinverse is $I^+ = \frac{|\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|}{\langle\!\langle \mathbb{1}|\mathbb{1}\rangle\!\rangle^2} = \frac{|\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|}{d^2}$. Eqs. (A9) and (A2) imply that

$$\mathcal{T}_\mathcal{I} = \big\{T = \alpha\mathbb{1} \,:\, \alpha \in \mathbb{C}, |\alpha| \leq 1\big\}. \tag{A10}$$

Any such $T = \alpha\mathbb{1}$ with $|\alpha| \leq 1$ can indeed be obtained by taking for instance $\{K_i\}_i = \{K_0 = \mathbb{1}\}$ and $\langle\varepsilon_0|0\rangle = \alpha$. As one can see, even the identity channel does not define a unique transformation matrix. The freedom one has on its possible transformation matrices is not just due to a possible global phase (which would just restrict α above to $|\alpha| = 1$), but also to the possible coherent control of some operation $|\psi\rangle^t \otimes |\varepsilon\rangle^e \to |\psi\rangle^t \otimes |0\rangle^e$ that (while acting trivially on the target system) acts nontrivially on the environment. Note that Eq. (A10) generalises straightforwardly to any unitary channel $\mathcal{U} : \rho \to U\rho U^\dagger$, whose possible transformation matrices are of the form $T = \alpha U$ with $|\alpha| \leq 1$.

For a d-dimensional fully depolarising channel $\mathcal{N}$, the Choi matrix is $N = \frac{1}{d}\mathbb{1}$; its range is the full Hilbert space $\mathcal{H}_I \otimes \mathcal{H}_O$, and its Moore-Penrose pseudoinverse is $N^+ = d\,\mathbb{1}$. Noting that $\langle\!\langle T|T\rangle\!\rangle = \mathrm{Tr}[T^\dagger T]$, Eq. (A9) implies that

$$\mathcal{T}_\mathcal{N} = \Big\{T \,:\, \mathrm{Tr}[T^\dagger T] \leq \frac{1}{d}\Big\}. \tag{A11}$$

Any such T satisfying $\mathrm{Tr}[T^\dagger T] \leq \frac{1}{d}$ can indeed be obtained by taking for instance the set of Kraus operators $\{\frac{1}{d}U_i\}_{i=0}^{d^2-1}$ (where the U_i's are again orthogonal unitary matrices) and $|\varepsilon\rangle^e$ such that $\langle\varepsilon_0|i\rangle = \langle\!\langle U_i|T\rangle\!\rangle = \mathrm{Tr}[U_i^\dagger T]$.

Combining the channels $\mathcal{I}$ and $\mathcal{N}$, the Choi matrix of a partially depolarising channel (as considered, e.g., in Ref. [3]) $\mathcal{N}_{(q)} \coloneqq q\,\mathcal{I} + (1-q)\,\mathcal{N}$, with $0 \leq q \leq 1$, is $N_{(q)} = q\,I + (1-q)\,N = q\,|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}| + (1-q)\,\frac{1}{d}\mathbb{1}$. For $q < 1$ its range is again the full Hilbert space $\mathcal{H}_I \otimes \mathcal{H}_O$; its Moore-Penrose pseudoinverse then coincides with its inverse, and is found to be $N^+_{(q)} = \frac{d}{1-q}[\mathbb{1} - \frac{dq}{d^2q+1-q}\,|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|]$. Noting that $\langle\!\langle\mathbb{1}|T\rangle\!\rangle = \mathrm{Tr}[T]$, Eq. (A9) then gives

$$\mathcal{T}_{\mathcal{N}_{(q)}} = \Big\{T \,:\, \mathrm{Tr}[T^\dagger T] - \frac{dq}{d^2q+1-q}\big|\,\mathrm{Tr}[T]\big|^2 \leq \frac{1-q}{d}\Big\}. \tag{A12}$$

Any T satisfying the above constraint can for instance be obtained with the Kraus operators $\{K_0 = \frac{\sqrt{d^2q+1-q}}{d}\mathbb{1}, K_i = \frac{\sqrt{1-q}}{d}U_i$ for $1 \leq i \leq d^2-1\}$ (using a set of orthogonal unitaries that contains $U_0 = \mathbb{1}$) and $\langle\varepsilon_0|0\rangle = \frac{1}{\sqrt{d^2q+1-q}}\,\mathrm{Tr}[T]$, $\langle\varepsilon_0|i\rangle = \frac{1}{\sqrt{1-q}}\,\mathrm{Tr}[U_i^\dagger T]$ for $1 \leq i \leq d^2-1$.

Consider, as another example, the partially dephasing qubit channel in the Pauli σ_z eigenbasis—or phase-flip channel—$\mathcal{Z}_{(p)}(\rho) \coloneqq (1-p)\,\rho + p\,\sigma_z\rho\sigma_z$ with $0 \leq p \leq 1$ (which is completely dephasing for $p = \frac{1}{2}$). Its Choi matrix representation is $Z_{(p)} = (1-p)\,|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}| + p\,|\sigma_z\rangle\!\rangle\langle\!\langle\sigma_z|$; for $0 < p < 1$ its range is the span of $\{|\mathbb{1}\rangle\!\rangle, |\sigma_z\rangle\!\rangle\}$ and its Moore-Penrose pseudoinverse is $Z^+_{(p)} = \frac{1}{4(1-p)}\,|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}| + \frac{1}{4p}\,|\sigma_z\rangle\!\rangle\langle\!\langle\sigma_z|$. Eq. (A9) leads to the characterisation (also valid for $p = 0$ or 1)

$$\mathcal{T}_{\mathcal{Z}_{(p)}} = \big\{T = \alpha\sqrt{1-p}\,\mathbb{1} + \beta\sqrt{p}\,\sigma_z : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 \leq 1\big\}. \tag{A13}$$

Any such $T = \alpha\sqrt{1-p}\,\mathbb{1} + \beta\sqrt{p}\,\sigma_z$ with $|\alpha|^2 + |\beta|^2 \leq 1$ can straightforwardly be obtained, for instance, from the Kraus operators $\{K_0 = \sqrt{1-p}\,\mathbb{1}, K_1 = \sqrt{p}\,\sigma_z\}$, by taking $\langle\varepsilon_0|0\rangle = \alpha$ and $\langle\varepsilon_0|1\rangle = \beta$.

Similarly, the set of possible transformation matrices for the partially dephasing qubit channel in the σ_x eigenbasis—or bit-flip channel—$\mathcal{X}_{(p)}(\rho) \coloneqq (1-p)\,\rho + p\,\sigma_x\rho\sigma_x$ is

$$\mathcal{T}_{\mathcal{X}_{(p)}} = \big\{T = \alpha\sqrt{1-p}\,\mathbb{1} + \beta\sqrt{p}\,\sigma_x : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 \leq 1\big\}. \tag{A14}$$

and any such $T = \alpha\sqrt{1-p}\,\mathbb{1} + \beta\sqrt{p}\,\sigma_x$ can be obtained from the Kraus operators $\{K_0 = \sqrt{1-p}\,\mathbb{1}, K_1 = \sqrt{p}\,\sigma_x\}$, with $\langle\varepsilon_0|0\rangle = \alpha$ and $\langle\varepsilon_0|1\rangle = \beta$.

B Communication of classical and quantum information through coherently controlled depolarising and dephasing channels

B.1 Holevo information of the coherently controlled depolarising channels

In this section we look at how much (classical) information can be transmitted by coherently controlling two depolarising channels (i.e., in the scenario of Fig. 2 with $\mathcal{C}_i = \mathcal{N}_i$). Recall that the Holevo information of a channel $\mathcal{C}$ quantifies how much classical information can be transmitted through a single use of $\mathcal{C}$ from a party A to another party B. It is defined as $\chi(\mathcal{C}) \coloneqq \max_{\{p_a,\rho_a\}} I(A;B)_\nu$, where $I(A;B)_\nu$ is the quantum mutual information calculated on the state $\nu \coloneqq \sum_a p_a\,|a\rangle\langle a|_A \otimes \mathcal{C}(\rho_a)_B$ [42, 43] (i.e., $I(A;B)_\nu = H(A)_\nu + H(B)_\nu - H(AB)_\nu$, where $H(X)_\nu$ is the von Neumann entropy of the system $X \in \{A, B, AB\}$ in the state ν). $\chi(\mathcal{C})$ provides a lower bound for the classical capacity of a quantum channel $\mathcal{C}$.

For a fixed initial state of the control qubit $\rho^c_{\text{in}} = |+\rangle\langle+|^c$, the depolarising quantum switch can be seen as a global channel, which we denote here $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$, mapping ρ^t_{in} to ρ^{ct}_{out} (see inset of Fig. 1 in main text), with $\rho^{ct}_{\text{out}} = \mathcal{S}[\mathcal{N}_0, \mathcal{N}_1](\rho^t_{\text{in}})$ given by Eq. (1) in the main text.[9] In Ref. [3] it was shown that the Holevo information of the depolarising quantum switch (with $\rho^c_{\text{in}} = |+\rangle\langle+|^c$) is $\chi(\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]) = -\frac{3}{8} - \frac{5}{8}\log_2\frac{5}{8} \simeq 0.05$ for a qubit target system, while a more general formula for any dimension d was also given.

The situation of Fig. 2 similarly induces a global channel, which we denote by $\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1]$, mapping ρ^t_{in} to ρ^{ct}_{out} according to Eq. (5) of the main text. A lower bound on the Holevo information $\chi(\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1])$ can (for a given pair of channels and transformation matrices) be easily obtained by simply taking $I(A;B)_\nu$, for any particular choice of the weighted ensemble $\{p_a, \rho_a\}$. For two fully depolarising channels $\mathcal{N}_0, \mathcal{N}_1$ with $T_0 = T_1 = \frac{1}{\sqrt{d}}\,|0\rangle\langle 0|$ for instance (which indeed satisfies $\mathrm{Tr}[T_i^\dagger T_i] \leq \frac{1}{d}$ as required by Eq. (A11)), taking $\{p_0 = \frac{3}{5}, \rho_0 = |0\rangle\langle 0|, p_1 = \frac{2}{5}, \rho_1 = |1\rangle\langle 1|\}$ gives the lower bound $\chi(\mathcal{M}[\mathcal{N}_0, T_0, \mathcal{N}_1, T_1]) \geq \frac{1}{d}\log_2\frac{5}{4}$, which is a significant increase over that obtained by the depolarising quantum switch. For $d = 2$ this indeed gives $\chi(\mathcal{M}[\mathcal{N}_0, T_0, \mathcal{N}_1, T_1]) \gtrsim 0.16$; for larger d this bound decreases as the dimension increases, but less rapidly than the (exact) Holevo information obtained in Ref. [3]. It remains an open question whether our lower bound is

[9] Note that as we consider a fixed initial state of the control qubit, which the sender cannot control, it is not part of the input of the channel $\rho^t_{\text{in}} \to \rho^{ct}_{\text{out}}$ considered here (but it contributes to the definition of the channel).

tight, both for the transformation matrices T_i as well as for any other transformation matrices for two completely depolarising channels.

One may be tempted to argue that the fact that the lower bound obtained here exceed the Holevo information for the depolarising quantum switch obtained in Ref. [3] is perhaps not so surprising given the differences between the scenario in Fig. 2 and that of the quantum switch. Indeed, in the scenario we consider, the target system only goes through the depolarising channels (in a superposition) a single time, while in the quantum switch the target system always goes through both channels (in a superposition of different orders). Thus, one may intuitively expect the target system to be less "degraded" in the scenario considered here. However, a recent paper exploring numerically this issue [39] (which appeared subsequent to an earlier preprint of this manuscript and, in part, in response to the very questions we raise) indicates that the situation is more subtle, with the noncommutativity of the Kraus operators of the two channels placed inside a quantum switch also playing a role in these communication advantages. In particular, in some cases the quantum switch indeed outperforms coherent control meaning this intuition—at least applied naively—is incorrect, although further study is needed to understand fully the situation.

B.2 Quantum information transfer through coherently controlled dephasing channels

Here we look at how much quantum information can be transmitted through two coherently controlled complementary dephasing channels. Recall that the quantum information that can be communicated through a channel $\mathcal{C}$ from some system A to some other system B is quantified by the quantum capacity $\mathcal{Q}(\mathcal{C})$ [57–59]. A lower bound on $\mathcal{Q}(\mathcal{C})$ is given by the coherent information $\mathcal{Q}^{(1)}(\mathcal{C}) \coloneqq \max_{\nu_0}[H(B)_\nu - H(AB)_\nu]$, where the maximisation is over all states $\nu_0 \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_{A'})$ of a bipartite system comprising A and a reference system A' isomorphic to A, and with $\nu \coloneqq (\mathcal{I} \otimes \mathcal{C})(\nu_0) \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ [42]. A further lower bound on $\mathcal{Q}^{(1)}(\mathcal{C})$ can of course be obtained by choosing any specific state ν_0.

Consider, as in Ref. [51], the case where the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ in Fig. 2 are (two-dimensional) phase-flip and the bit-flip channels $\mathcal{Z}_{(p)}$ and $\mathcal{X}_{(p)}$, respectively (for simplicity we take the same mixing parameter p for both channels). The possible transformation matrices for these channels are characterised by Eqs. (A13) and (A14). Taking $T_{\mathcal{Z}_{(p)}} = \sqrt{p}\,\sigma_z$ and $T_{\mathcal{X}_{(p)}} = \sqrt{p}\,\sigma_x$ (i.e., taking $\alpha = 0, \beta = 1$ in Eqs. (A13)–(A14)), and writing $\mathcal{M}_{(\mathcal{Z},\mathcal{X},p)} \coloneqq \mathcal{M}[\mathcal{Z}_{(p)}, T_{\mathcal{Z}_{(p)}}, \mathcal{X}_{(p)}, T_{\mathcal{X}_{(p)}}]$, we find, from Eq. (5) of the main text, that

$$\mathcal{M}_{(\mathcal{Z},\mathcal{X},p)}(\rho^t_{\text{in}}) = (1-p)\frac{\mathbb{1}}{2} \otimes \rho^t_{\text{in}} + p\,\frac{|0\rangle^c \otimes \sigma_z + |1\rangle^c \otimes \sigma_x}{\sqrt{2}}\rho^t_{\text{in}}\frac{\langle 0|^c \otimes \sigma_z + \langle 1|^c \otimes \sigma_x}{\sqrt{2}}. \tag{B1}$$

This allows one to calculate $\nu \coloneqq (\mathcal{I} \otimes \mathcal{M}_{(\mathcal{Z},\mathcal{X},p)})(\nu_0)$, and then $H(B)_\nu - H(AB)_\nu$, for any choice of ν_0. By taking the maximally entangled state $\nu_0 = |\Phi^+\rangle\langle\Phi^+|$ with $|\Phi^+\rangle \coloneqq \frac{1}{\sqrt{2}}(|0\rangle\otimes|0\rangle + |1\rangle\otimes|1\rangle)$, after some calculation we obtain

$$\mathcal{Q}(\mathcal{M}_{(\mathcal{Z},\mathcal{X},p)})) \geq \mathcal{Q}^{(1)}(\mathcal{M}_{(\mathcal{Z},\mathcal{X},p)})) \geq p - H_2(p) + H_2\big(\frac{1-p}{2}\big), \tag{B2}$$

where H_2 is the binary entropy function $H_2(p) \coloneqq -p\,\log_2(p) - (1-p)\log_2(1-p)$.

This lower bound on the quantum capacity is larger than that obtained in Ref. [51] for the full quantum switch with phase-flip and bit-flip channels. Remarkably, our bound is positive for all values of p—in particular, it takes the value $-\frac{3}{4}\log_2\frac{3}{4} \simeq 0.31$ for $p = \frac{1}{2}$, i.e., for fully dephasing channels (which, by themselves, cannot transmit any quantum information). It is also larger than the quantum capacities of each channel $\mathcal{Z}_{(p)}$ and $\mathcal{X}_{(p)}$ individually—and thus violates the "bottleneck inequality" considered in Ref. [51]—for all p. In comparison, the bound obtained in Ref. [51] with the full quantum switch was positive only for $p \lesssim 0.13$ and $p \gtrsim 0.60$, and was larger than $\mathcal{Q}(\mathcal{Z}_{(p)})$ or $\mathcal{Q}(\mathcal{X}_{(p)})$ only for $p \gtrsim 0.62$.

As for the Holevo information, we remark that, in our scenario, the target system only ever passes through a single dephasing channel, which may, at least in part, help explain the higher bound we obtain on the quantum capacity. Once again, however, the situation merits further study, as shown by the results of Ref. [52] mentioned in Sec. 6 which show that the quantum switch can, in some cases, enable perfect quantum communication where coherent control cannot.

C Channel implementation independence for the full quantum switch and other quantum processes

For the case of the full quantum switch, each of the two channels $\mathcal{C}_0$ and $\mathcal{C}_1$, with Kraus operators $\{K_i\}_i$ and $\{L_j\}_j$, is applied once and only once on the target system. Considering a purified version of the channels via a

Stinespring dilation, as described in the main text, the state at the output of the interferometer (see the inset of Fig. 1) reads

$$\frac{1}{\sqrt{2}}\left|0\right\rangle^c \otimes \sum_{i,j} L_j K_i \left|\psi_{\text{in}}\right\rangle^t \otimes \left|i\right\rangle^{e_0} \otimes \left|j\right\rangle^{e_1} + \frac{1}{\sqrt{2}}\left|1\right\rangle^c \otimes \sum_{i,j} K_i L_j \left|\psi_{\text{in}}\right\rangle^t \otimes \left|i\right\rangle^{e_0} \otimes \left|j\right\rangle^{e_1}. \tag{C1}$$

In contrast to the output state (5) for the circuit of Fig. 1, no terms appear in which either environment is untouched and remains in its initial state. After tracing out the environments, one obtains

$$\rho_{\text{out}}^{ct} = \frac{1}{2}\Big(\left|0\right\rangle\!\left\langle 0\right|^c \otimes \mathcal{C}_1 \circ \mathcal{C}_0(\rho_{\text{in}}^t) + \left|1\right\rangle\!\left\langle 1\right|^c \otimes \mathcal{C}_0 \circ \mathcal{C}_1(\rho_{\text{in}}^t) + \left|0\right\rangle\!\left\langle 1\right|^c \otimes \sum_{i,j} L_j K_i \rho_{\text{in}}^t L_j^\dagger K_i^\dagger + \left|1\right\rangle\!\left\langle 0\right|^c \otimes \sum_{i,j} K_i L_j \rho_{\text{in}}^t K_i^\dagger L_j^\dagger \Big), \tag{C2}$$

which depends neither on the initial state of the environments, nor on the sets of Kraus operators chosen to describe each channel. Indeed, for any other Kraus representations $\{M_r\}_r$ of $\mathcal{C}_0$ and $\{N_s\}_s$ of $\mathcal{C}_1$, one has $K_i = \sum_r u_{ir} M_r$ and $L_j = \sum_s v_{js} N_s$, where u_{ir} and v_{js} are the elements of unitary matrices [1]. We thus obtain

$$\begin{aligned} \sum_{i,j} L_j K_i \rho_{\text{in}}^t L_j^\dagger K_i^\dagger &= \sum_{i,j} \sum_{r,r',s,s'} u_{ir} u_{ir'}^* v_{js} v_{js'}^* N_s M_r \rho_{\text{in}}^t N_{s'}^\dagger M_{r'}^\dagger \\ &= \sum_{r,r',s,s'} \delta_{r,r'} \delta_{s,s'} N_s M_r \rho_{\text{in}}^t N_{s'}^\dagger M_{r'}^\dagger = \sum_{r,s} N_s M_r \rho_{\text{in}}^t N_s^\dagger M_r^\dagger \end{aligned} \tag{C3}$$

(where δ is the Kronecker delta), and analogously for the term $\sum_{i,j} K_i L_j \rho_{\text{in}}^t K_i^\dagger L_j^\dagger$.

More generally, consider a combination of multiple channels $\mathcal{C}_0, \ldots, \mathcal{C}_N$ with Kraus operators $\{K_{i_0}^{(0)}\}_{i_0}, \ldots, \{K_{i_N}^{(N)}\}_{i_N}$, and assume that for any possible initial state $\left|\Psi_{\text{in}}\right\rangle$ sent through the setup, each channel is applied once and only once (not necessarily in a definite order). Considering a Stinespring dilation of the channels with environment initial states $\left|\varepsilon_0\right\rangle^{e_0}, \ldots, \left|\varepsilon_N\right\rangle^{e_N}$, this means that the joint initial state evolves as

$$\left|\Psi_{\text{in}}\right\rangle \otimes \left|\varepsilon_0\right\rangle^{e_0} \otimes \cdots \otimes \left|\varepsilon_N\right\rangle^{e_N} \to \sum_{i_0,\ldots,i_N} F(K_{i_0}^{(0)}, \ldots, K_{i_N}^{(N)}) \left|\Psi_{\text{in}}\right\rangle \otimes \left|i_0\right\rangle^{e_0} \otimes \cdots \otimes \left|i_N\right\rangle^{e_N}, \tag{C4}$$

where each $F(K_{i_0}^{(0)}, \ldots, K_{i_N}^{(N)})$ is an operator composed as a sum of product terms in which each $K_{i_\ell}^{(\ell)}$ appears once and only once, in possibly different orders (e.g., for the quantum switch: $\left|\Psi_{\text{in}}\right\rangle = \left|+\right\rangle^c \otimes \left|\psi_{\text{in}}\right\rangle^t$ and $F(K_i, L_j) = \left|0\right\rangle\!\left\langle 0\right|^c \otimes L_j K_i + \left|1\right\rangle\!\left\langle 1\right|^c \otimes K_i L_j$). For any such transformation, a similar calculation as for the full quantum switch can be conducted, which shows that after tracing out the environments, the final output state does not depend on the choice of Kraus operators, nor on the initial states of the environments.

We note that the assumption that each channel is applied—or that "each party acts"—once and only once is at the core of the process matrix framework [12, 60]. This justifies, beyond the particular case of the quantum switch, that the situations described by process matrices do not depend on any specific implementation of the channels (or of more general quantum operations) applied by each party, but only on the description (e.g., in terms of Kraus operators) of the induced completely positive maps.

D Distinguishing different implementations of coherently-controlled channels

In this section we show that the diamond-norm distance in the scenario of Fig. 1, where the channel $\mathcal{C}_0$ has a fixed transformation matrix T_0 while the channel $\mathcal{C}_1$ has two possible transformation matrices T_1 and T_1', is obtained and bounded as in Eq. (6) of the main text.

Recall that the diamond-norm distance between the two channels $\mathcal{M}_{T_1} \coloneqq \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1]$ and $\mathcal{M}_{T_1'} \coloneqq \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1']$ is defined as

$$\mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}) \coloneqq \frac{1}{2}\left\|\mathcal{M}_{T_1} - \mathcal{M}_{T_1'}\right\|_\diamond = \frac{1}{2}\max_{\rho_{\text{in}}^{tr}} \left\|(\mathcal{M}_{T_1} \otimes \mathcal{I})(\rho_{\text{in}}^{tr}) - (\mathcal{M}_{T_1'} \otimes \mathcal{I})(\rho_{\text{in}}^{tr})\right\|_1, \tag{D1}$$

where $\|\cdot\|_1$ is the trace norm and the maximum is taken over all joint states ρ_{in}^{tr} of the target system (t) and some arbitrary reference system (r). That is, it is the maximum trace-distance between output states when the channels are applied to part of joint system (and with the identity channel $\mathcal{I}$ being applied to the other part); indeed, considering an input state entangled with a reference system is in general necessary for optimal channel discrimination [50].

Eq. (5) can straightforwardly be adapted to give the corresponding output state ρ_{out}^{ctr} of the control, target and reference systems, by replacing ρ_{in}^{t} by ρ_{in}^{tr}, $\mathcal{C}_{0/1}$ by $\tilde{\mathcal{C}}_{0/1} \coloneqq \mathcal{C}_{0/1} \otimes \mathcal{I}$, and $T_{0/1}$ by $\tilde{T}_{0/1} \coloneqq T_{0/1} \otimes \mathbb{1}$. Defining $\tau_1 \coloneqq T_1 - T_1'$ and $\tilde{\tau}_1 \coloneqq \tau_1 \otimes \mathbb{1} = \tilde{T}_1 - \tilde{T}_1'$ for ease of notation, we have, for an arbitrary input state ρ_{in}^{tr},

$$\begin{aligned}\frac{1}{2}\left\|(\mathcal{M}_{T_1} \otimes \mathcal{I})(\rho_{\text{in}}^{tr}) - (\mathcal{M}_{T_1'} \otimes \mathcal{I})(\rho_{\text{in}}^{tr})\right\|_1 &= \frac{1}{2}\operatorname{Tr}\sqrt{(\rho_{\text{out}}^{ctr} - \rho_{\text{out}}^{ctr\,\prime})^2} \\ &= \frac{1}{4}\operatorname{Tr}\sqrt{|0\rangle\langle 0|^c \otimes \tilde{T}_0 \rho_{\text{in}}^{tr} \tilde{\tau}_1^\dagger \tilde{\tau}_1 \rho_{\text{in}}^{tr} \tilde{T}_0^\dagger + |1\rangle\langle 1|^c \otimes \tilde{\tau}_1 \rho_{\text{in}}^{tr} \tilde{T}_0^\dagger \tilde{T}_0 \rho_{\text{in}}^{tr} \tilde{\tau}_1^\dagger} \\ &= \frac{1}{4}\left(\left\|\tilde{\tau}_1 \rho_{\text{in}}^{tr} \tilde{T}_0^\dagger\right\|_1 + \left\|\tilde{T}_0 \rho_{\text{in}}^{tr} \tilde{\tau}_1^\dagger\right\|_1\right) = \frac{1}{2}\left\|\tilde{\tau}_1 \rho_{\text{in}}^{tr} \tilde{T}_0^\dagger\right\|_1, \end{aligned} \tag{D2}$$

where the last equality follows from the invariance of the trace norm under Hermitian transposition.

Decomposing ρ_{in}^{tr} in the form $\rho_{\text{in}}^{tr} = \sum_m p_m |\psi_m\rangle\langle\psi_m|$ (with $p_m \geq 0$, $\sum_m p_m = 1$, $\langle\psi_m|\psi_m\rangle = 1\ \forall m$), using the sub-additivity of the trace norm, and the fact that $\||x\rangle\langle y|\|_1 = \sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle}$ for any two vectors $|x\rangle$, $|y\rangle$, we further have

$$\frac{1}{2}\left\|\tilde{\tau}_1 \rho_{\text{in}}^{tr} \tilde{T}_0^\dagger\right\|_1 \leq \frac{1}{2}\sum_m p_m \left\|\tilde{\tau}_1 |\psi_m\rangle\langle\psi_m| \tilde{T}_0^\dagger\right\|_1 = \frac{1}{2}\sum_m p_m \sqrt{\langle\psi_m| \tilde{\tau}_1^\dagger \tilde{\tau}_1 |\psi_m\rangle}\sqrt{\langle\psi_m| \tilde{T}_0^\dagger \tilde{T}_0 |\psi_m\rangle}. \tag{D3}$$

Now, by the definition of the spectral norm one has $\sqrt{\langle\psi_m| \tilde{\tau}_1^\dagger \tilde{\tau}_1 |\psi_m\rangle} \leq \|\tilde{\tau}_1\|_2 = \|\tau_1\|_2$ for all m. Additionally,

$$\langle\psi_m| \tilde{T}_0^\dagger \tilde{T}_0 |\psi_m\rangle = \Big(\sum_{i'} \big(\langle\psi_m| K_{i'}^\dagger \otimes \mathbb{1}^r\big) \otimes \langle i'|^{e_0}\Big)\Big(\mathbb{1}^t \otimes \mathbb{1}^r \otimes |\epsilon_0\rangle\langle\epsilon_0|^{e_0}\Big)\Big(\sum_i \big(K_i \otimes \mathbb{1}^r |\psi_m\rangle\big) \otimes |i\rangle^{e_0}\Big) \leq 1. \tag{D4}$$

Combined together, these bounds lead directly to the upper bound of Eq. (6). Furthermore, it is easy to see that the inequalities above—and thus the upper bound of Eq. (6)—can always be saturated by taking $T_0 = \mathbb{1}$ (with $\mathcal{C}_0 = \mathcal{I}$) and $\rho_{\text{in}}^{tr} = \rho_{\text{in}}^{t} = |\psi\rangle\langle\psi|$ with $|\psi\rangle$ maximising $\langle\psi| \tau_1^\dagger \tau_1 |\psi\rangle$, i.e., without any (or with a trivial) reference system.

To finish with, we note that in the example presented in the main text in which $\mathcal{C}_1 = \mathcal{N}$, the two implementations we proposed, with $T_1, T_1' = \pm\frac{1}{\sqrt{d}}|0\rangle\langle 0|$, are the most distinguishable ones that can be considered. Indeed, according to Eq. (A11) any transformation matrix T_1 of $\mathcal{C}_1 = \mathcal{N}$ must satisfy $\operatorname{Tr}[T_1^\dagger T_1] = \left\|T_1\right\|_{\text{HS}}^2 \leq \frac{1}{d}$ (with $\|\cdot\|_{\text{HS}}$ denoting the Hilbert-Schmidt norm; this constraint is indeed satisfied for our choice above). Using the fact that the Hilbert-Schmidt norm always upper-bounds the spectral norm, it follows that

$$\frac{1}{2}\left\|T_1 - T_1'\right\|_2 \leq \frac{1}{2}\left\|T_1 - T_1'\right\|_{\text{HS}} \leq \frac{1}{2}\left(\left\|T_1\right\|_{\text{HS}} + \left\|T_1'\right\|_{\text{HS}}\right) \leq \frac{1}{\sqrt{d}}, \tag{D5}$$

so that the value one obtains with the choice above, $\frac{1}{2}\|T_1 - T_1'\|_2 = \frac{1}{\sqrt{d}}$, is indeed the largest possible one.