Witnessing latent time correlations with a single quantum particle

Hlér Kristjánsson [1,2] Wenxu Mao (貌文徐),[3,4] and Giulio Chiribella [3,1,2,5,*]

[1]*Quantum Group, Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom*
[2]*HKU-Oxford Joint Laboratory for Quantum Information and Computation*
[3]*QICI Quantum Information and Computation Initiative, Department of Computer Science,*
The University of Hong Kong, Pokfulam Road, Hong Kong
[4]*Cornell Tech, Cornell University, 2 West Loop Road, New York, New York 10044, USA*
[5]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5*

When a noisy communication channel is used multiple times, the errors occurring at different times generally exhibit correlations. Classically, these correlations do not affect the evolution of individual particles: a single classical particle can only traverse the channel at a definite moment of time, and its evolution is insensitive to the correlations between subsequent uses of the channel. In stark contrast, here we show that a single quantum particle can sense the correlations between multiple uses of a channel at different moments of time. Taking advantage of this phenomenon, it is possible to enhance the amount of information that the particle can reliably carry through the channel. In an extreme example, we show that a channel that outputs white noise whenever the particle is sent at a definite time can exhibit correlations that enable a perfect transmission of classical bits when the particle is sent at a superposition of two distinct times. In contrast, we show that, in the lack of correlations, a single particle sent at a superposition of two times undergoes an effective channel with classical capacity of at most 0.16 bits. When multiple transmission lines are available, time correlations can be used to simulate the application of quantum channels in a coherent superposition of alternative causal orders, and even to provide communication advantages that are not accessible through the superposition of causal orders.

I. INTRODUCTION

Quantum communication enables new possibilities that were unthinkable in the classical world, notably including secure key distribution [1,2]. The main hurdle to the implementation of quantum communication, however, is the fragility of quantum states to noise. To tackle this problem, quantum error correction schemes encode information into multiple quantum particles, using redundancy to mitigate the effects of noise [3–5].

When the same communication channel is used multiple times, the noisy processes experienced by particles sent at different times are generally correlated [6–9]. For example, photons transmitted through an optical fiber are subject to random changes in their polarization [10], and since such changes happen on a finite timescale, photons sent at nearby times experience approximately the same noisy processes. A similar situation arises in satellite quantum communication, where the satellite's motion induces dynamical mismatches of reference frame with respect to the ground station [11].

The presence of correlations is both a threat and an opportunity for communication. On the one hand, it can undermine the effectiveness of standard error-correcting schemes, which assume independent errors on the transmitted particles. On the other hand, tailored codes that exploit the correlations among different particles can enhance the transmission of information [6,8,12–26].

Like most error-correcting schemes, the existing codes for correlated noise use multiple physical particles to encode a single logical message. Classically, the use of multiple particles is essential: since a single classical particle can only traverse a communication channel at a definite moment of time, correlations between different uses of the channel do not affect the particle's evolution. The same conclusion holds even if the moment of transmission is chosen at random: in this case, the resulting evolution is simply the average of the evolutions associated to each individual moment of time, and the overall evolution is independent of the time correlations.

In stark contrast, here we show that a single quantum particle can sense the correlations between multiple uses of the same quantum communication channel. At the fundamental level, this effect is made possible by the ability of quantum particles to experience a coherent superposition of multiple time evolutions [27–34]. In particular, we will consider the situation in which the particle is in a superposition of traveling at different moments of time, as illustrated in Fig. 1. Taking advantage of the time correlations in the noise, we show that it is possible to enhance the amount of information that a
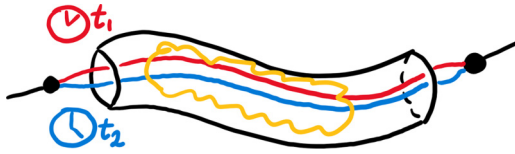
---

*giulio.chiribella@cs.ox.ac.uk

FIG. 1. A single quantum particle can travel through a transmission line at a superposition of two different moments of time t_1 (red) and t_2 (blue). Along the way, the particle experiences errors (yellow region), and the errors occurring at time t_1 are generally correlated with the errors occurring at time t_2. By taking advantage of these correlations, the errors can be mitigated or even completely removed.

single particle can carry from a sender to a receiver, beating the ultimate limit achievable in the lack of correlations.

We demonstrate this effect with an extreme example, in which a single quantum particle carries one bit of classical information through a transmission line that completely erases information at every definite time step. This phenomenon witnesses the presence of correlations between different uses of the transmission line: in the lack of correlations, we show that the number of bits that can be reliably transmitted by sending a single particle at a superposition of two different times does not exceed 0.16.

It is worth stressing that the above advantage is not specific to time correlations, but applies more generally to spatial correlations, or to other types of correlations: as long as two different uses of a channel are correlated, one may take advantage of the correlations by sending a quantum particle in a superposition of going through one use or the other.

Time-correlated channels are also interesting for foundational reasons. Recently, they have been proposed as a way to reproduce the use of quantum channels in a superposition of different causal orders [32,35]. In particular, they have been used to reproduce the action of the quantum SWITCH [36,37], a higher-order operation that combines two variable quantum channels in a superposition of two alternative orders. In practice, time-correlated channels underlie all the existing experimental setups inspired by the quantum SWITCH [38–44].

The quantum SWITCH is known to offer a number of advantages in quantum communication [42,45–50]. Here we show that (1) time correlations are essential in order to reproduce the advantages of the quantum SWITCH, and (2) the access to time-correlated channels is an even more powerful resource than the ability to combine ordinary quantum channels in a superposition of alternative orders.

To make the above points, we consider the scenario illustrated in Fig. 2, where a single particle is sent on a superposition of two paths, traversing two independent channels, each with the property that different uses of the same channel at different moments of time are correlated, while the action of the channel at any given time is completely depolarizing. When the noise is perfectly correlated, the network in Fig. 2 reproduces the quantum SWITCH of two completely depolarizing channels, which is known to achieve a communication capacity of 0.049 [45,50]. In contrast, we show that in the lack of time correlations the maximum capacity achieved by sending a particle on a superposition of paths is at most 0.024 bits. This result proves that, in this scenario, the physical origin of the communication advantage of the
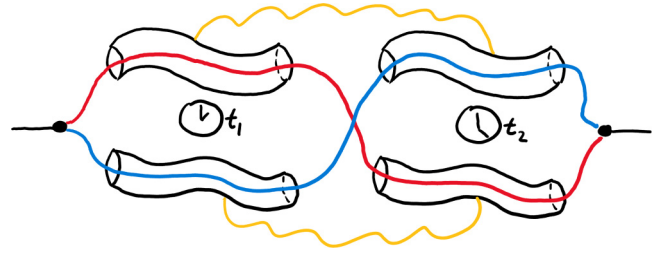

FIG. 2. A single particle can travel on a superposition of two different paths (red and blue), which traverse two transmission lines (top and bottom) at two moments of time t_1 and t_2. The errors occurring on successive uses of the same transmission line are correlated (yellow lines), so the particle experiences correlated errors across the two branches (red and blue) of the superposition. These time correlations are a resource that can be used to mimic the use of quantum channels in a superposition of orders, and even to achieve larger communication advantages.

quantum SWITCH is not merely the superposition of paths, but rather the interplay between the superposition of paths and the time correlations in the noise.

Remarkably, we also find that the time correlations that reproduce the action of the quantum SWITCH are not the most favorable for the transmission of classical information: while the quantum SWITCH of two completely depolarizing channels can at most yield 0.049 bits of classical communication [45,50], a more sophisticated pattern of time correlations yields at least 0.31 bits. The gap between these two values further highlights the power of time correlations, which are not only capable of reproducing the benefits of the superposition of causal orders, but also of surpassing them.

The remainder of the paper is structured as follows. In Sec. II we describe the formalism of time-correlated channels and derive the effective evolution experienced by a single particle upon entering a time-correlated channel at a superposition of times. In Sec. III, we consider the transmission of a single particle at a superposition of times, as in Fig. 1, and we demonstrate that the correlations between different uses of the channels offer a communication advantage over all communication scenarios where the channels are uncorrelated. In Sec. IV, we consider the network scenario of Fig. 2, and we show that time correlations are necessary to reproduce the advantages of the quantum SWITCH, and that certain time correlations can even offer higher advantages. Finally, we discuss the effects of noise on the control degree of freedom in Sec. V and conclude in Sec. VI.

II. TRANSMISSION OF A SINGLE PARTICLE AT A SUPERPOSITION OF DIFFERENT TIMES

A. Time-correlated channels

A transmission line that can be accessed at k different times is described by a correlated quantum channel [6–8]. Mathematically, the correlated channel is a linear map transforming density matrices of the composite system $S_1 \otimes \cdots \otimes S_k$, where S_j denotes the system sent at the jth time. Note that, in general, the k systems sent at k different times can be initially prepared in an arbitrary entangled state.

Correlated quantum channels are also known as quantum memory channels [7,8,20], quantum combs [51,52], or non-Markovian quantum processes [9,53]. In the following we will focus on the $k = 2$ case, corresponding to a transmission line that can be accessed at two different time steps, hereafter denoted by t_1 and t_2. We consider random unitary channels of the form

$$\mathcal{R}(\rho_{12}) = \sum_{m,n} p(m,n)\,(U_m \otimes U_n)\,\rho_{12}\,(U_m \otimes U_n)^\dagger, \quad (1)$$

where U_m and U_n are unitary gates in a given set, and $p(m,n)$ is a joint probability distribution. Here, the system sent at time t_1 experiences the unitary gate U_m, while the system sent at time t_2 experiences the gate U_n. The density matrix ρ_{12} represents the joint state of the two systems sent at the two times t_1 and t_2, that is, ρ_{12} is a density matrix on the Hilbert space of the composite system $S_1 \otimes S_2$. The probability distribution $p(m,n)$ specifies the correlations between the random unitary evolutions experienced by system S_1 and system S_2.

Note that, while in this paper we will focus on time correlations, the correlations in Eq. (1) are not specific to time. The same expression can be used also to describe correlated channels acting on two spatially separated systems, or on any other type of independently addressable systems.

Physically, a time-correlated random unitary channel of the form (1) can arise in a photonic setup where the systems S_1 and S_2 are modes of the electromagnetic field associated to two different time bins [54–57]. The noisy channel can correspond, e.g., to the action of an optical fiber, where the random unitary changes of the photon polarization arise from random fluctuations in the birefringence. Correlations between the unitaries at different times can arise when the time difference $t_2 - t_1$ between successive uses of the channels is smaller than the timescale on which the birefringence fluctuates.

B. Sending a single particle through a time-correlated channel

Consider now the situation where the input of the correlated channel (1) is a single particle, carrying information in its internal degrees of freedom. Classically, the particle must be sent either at time t_1, or at time t_2, or at some random mixture of t_1 and t_2. When the particle is sent at time t_1, its evolution is given by the reduced channel $\mathcal{R}_1(\rho) := \sum_m p_1(m)\,U_m \rho U_m^\dagger$, where $p_1(m) := \sum_n p(m,n)$ is the marginal probability distribution of the unitaries at time t_1. Similarly, if the particle is sent at time t_2, its evolution is given by the channel $\mathcal{R}_2(\rho) := \sum_m p_2(n)\,U_n \rho U_n^\dagger$, with $p_2(n) := \sum_m p(m,n)$. A random choice of transmission times then results into a random mixture of the evolutions corresponding to channels $\mathcal{R}_1$ and $\mathcal{R}_2$. Crucially, the evolution of the particle is independent of any correlation that may be present in the probability distributions $p(m,n)$, that is, of any correlation between the first and the second use of the transmission line.

In contrast, quantum mechanics allows one to transmit a single particle in a way that is sensitive to the correlations between noisy processes at different times. The key idea is that the time when the particle is transmitted can be indefinite, as the particle could be sent through the transmission line at a coherent superposition of times t_1 and t_2 (see illustration in Fig. 3). The superposition of transmission times could
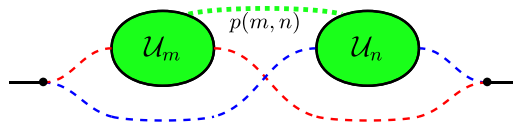

FIG. 3. A single particle is sent at a superposition of two times (red and blue dashed lines), through the same transmission line (green ovals). The green dotted line represents the correlations between random unitary processes $\mathcal{U}_m$ and $\mathcal{U}_n$ taking place with probability $p(m,n)$ at the two subsequent uses of the transmission line, respectively.

be achieved by adding an interferometric setup before the transmission line, letting the particle travel on a coherent superposition of two paths, one of which includes a delay [58]. This results in a time-bin qubit, described by a superposition of amplitudes corresponding to localization at two different points in time, separated by a time difference much greater than a photon's coherence time [59].

Before developing the general theory of single-particle transmission through time-correlated channels, it is instructive to look at a concrete example. Consider the case of a single photon, and denote by H_1 and V_1 (H_2 and V_2) the horizontal and vertical polarization modes in the first (second) time bin. Here we take the polarization state to be the same on both paths, so that the only role of the interferometric setup is to coherently control the moment of transmission. The result is a linear combination of states of the form $(\alpha|1\rangle_{H1}|0\rangle_{V1} + \beta|0\rangle_{H1}|1\rangle_{V1}) \otimes |0\rangle_{H2}|0\rangle_{V2}$ and states of the form $|0\rangle_{H1}|0\rangle_{V1} \otimes (\alpha|1\rangle_{H2}|0\rangle_{V2} + \beta|0\rangle_{H2}|1\rangle_{V2})$. The composite system of the two modes in the first (second) time bin can be regarded as system S_1 (S_2) in Eq. (1). The states produced by the interferometric setup can then be written as a linear combination of states of the form $|\psi\rangle_1 \otimes |\text{vac}\rangle_2$ and states of the form $|\text{vac}\rangle_1 \otimes |\psi\rangle_2$, where, for $i \in \{1,2\}$, $|\text{vac}\rangle_i := |0\rangle_{Hi}|0\rangle_{Vi}$ is the vacuum state of the modes in system S_i, and $|\psi\rangle_i := \alpha|1\rangle_{Hi}|0\rangle_{Vi} + \beta|0\rangle_{Hi}|1\rangle_{Vi}$ is a single-photon polarization state. The change in the particle's state upon the transmission is then computed by applying the channel (1) to the appropriate state.

Generalizing the above example, we model the transmission of a single particle through channel (1) by interpreting systems S_1 and S_2 as *abstract modes*, each of which can contain a variable number of particles equipped with an internal degree of freedom, such as the photon's polarization. For $i \in \{1,2\}$, the Hilbert space of system S_i has two orthogonal subspaces: a one-particle subspace, denoted by $A^{(i)}$, and a vacuum subspace, denoted by $\text{Vac}^{(i)}$. We assume that the dimension of the one-particle subspace is the same for both S_1 and S_2, as in the example of the single-photon polarization. Under this assumption, we have $A^{(1)} \simeq A^{(2)} \simeq M$, where M is the internal degree of freedom of the particle. Also, we assume that each vacuum subspace is one dimensional, and is spanned by a vacuum state $|\text{vac}\rangle_i$, $i \in \{1,2\}$, as in our motivating example.

A single particle sent at a superposition of two moments of time will then be described by states of the form $\alpha|\psi\rangle_1 \otimes |\text{vac}\rangle_2 + \beta|\text{vac}\rangle_1 \otimes |\psi\rangle_2$, where $|\psi\rangle \in M$ is the state of the particle's internal degree of freedom. For the transmission of the particle, we will consider channels that conserve the

number of particles, i.e., that map states of a given sector into states of the same sector. This is the case, for example, for linear optical elements, which preserve the photon number. For the channel (1), preservation of the particle number means that the operators U_m have the form

$$U_m = V_m + e^{i\phi_m} |\text{vac}\rangle\langle\text{vac}|, \qquad (2)$$

where V_m is a unitary acting in the one-particle sector M, and $\phi_m \in [0, 2\pi)$ is a phase. Physically, ϕ_m corresponds to the phase difference between states in the one-particle sector and the vacuum state.

In the quantum optical example, each unitary U_m can be realized by a Hamiltonian acting on the two polarization modes associated to system S_i, $i \in \{1, 2\}$. For example, the unitary $Z \oplus e^{i\phi} |\text{vac}\rangle\langle\text{vac}|$ can be generated by the Hamiltonian $H = \hbar[(\xi + \theta/2)a_H^\dagger a_H + (\xi - \theta/2)a_V^\dagger a_V]$, where a_H (a_V) are the annihilation operators for the appropriate modes with horizontal (vertical) polarization, in suitable units.

C. Effective evolution with a control system

The representation of a single particle in terms of abstract modes is equivalent to a representation in terms of a composite system MC, consisting of a message-carrying system M and a control system C, which determines the particle's time of transmission. The change of representation is described by the mapping

$$|\psi\rangle_1 \otimes |\text{vac}\rangle_2 \longmapsto |\psi\rangle_M \otimes |0\rangle_C$$
$$|\text{vac}\rangle_1 \otimes |\psi\rangle_2 \longmapsto |\psi\rangle_M \otimes |1\rangle_C, \qquad (3)$$

where $|\psi\rangle$ is an arbitrary state in the one-particle subspace. If the control is in state $|0\rangle$, then the message is sent through the first application of the channel, with the vacuum in the second application; *vice versa* if the control is in state $|1\rangle$. If the control is in a generic state ω, the overall evolution is described by an *effective channel* $\mathcal{C}_\omega$, which transforms a generic state ρ of the message into the state

$$\mathcal{C}_\omega(\rho) := \sum_{m,n} p(m, n) W_{mn} (\rho \otimes \omega) W_{mn}^\dagger, \qquad (4)$$

where W_{mn} is the unitary $W_{mn} := V_m e^{i\phi_n} \otimes |0\rangle\langle0| + e^{i\phi_m} V_n \otimes |1\rangle\langle1|$. The derivation of Eq. (4) is provided in Appendix A.

When the probability distribution $p(m, n)$ is symmetric [that is, when $p(m, n) = p(n, m)$ for every m and n], the effective channel has the simple expression

$$\mathcal{C}_\omega(\rho) = \frac{\mathcal{C}(\rho) + \mathcal{G}(\rho)}{2} \otimes \omega + \frac{\mathcal{C}(\rho) - \mathcal{G}(\rho)}{2} \otimes Z\omega Z, \quad (5)$$

with

$$\mathcal{C}(\rho) := \sum_{m,n} p(m, n) V_m \rho V_m^\dagger \qquad (6)$$

and

$$\mathcal{G}(\rho) := \sum_{m,n} p(m, n) e^{i(\phi_n - \phi_m)} V_m \rho V_n^\dagger. \qquad (7)$$

(See Appendix A for the derivation.) Here, the map $\mathcal{C}$ is the quantum channel representing the evolution of the message when it is sent at a definite time (either t_1 or t_2). The

channel $\mathcal{C}$ depends only on the marginal probability distribution $p_1(m) := \sum_n p(m, n)$, and it is independent of the correlations. Instead, the map $\mathcal{G}$ can generally depend on the correlations between the evolution of the particle at two mutually exclusive moments of time. We call $\mathcal{G}$ the *interference term*.

III. CLASSICAL COMMUNICATION THROUGH CORRELATED WHITE NOISE

A. Correlated white noise

Consider the case where the evolution at any definite time step is completely depolarizing on the message-carrying sector M, that is,

$$\mathcal{C}_{|j\rangle\langle j|}(\rho) = \frac{I}{d} \otimes |j\rangle\langle j|, \quad \forall \rho, \forall j \in \{0, 1\} \qquad (8)$$

where $\mathcal{C}_{|j\rangle\langle j|}$ is the quantum channel obtained by plugging $\omega = |j\rangle\langle j|$ into Eq. (4). Equation (8) implies that, whenever the particle is sent at a definite moment of time, the message is replaced by white noise. Accordingly, the channel $\mathcal{C}$ in Eq. (6) is depolarizing.

When the probability distribution $p(m, n)$ is symmetric, Eq. (5) becomes

$$\mathcal{C}_\omega(\rho) = \frac{I/d + \mathcal{G}(\rho)}{2} \otimes \omega + \frac{I/d - \mathcal{G}(\rho)}{2} \otimes Z\omega Z. \qquad (9)$$

In the realization of the random unitary channel, we will take the unitaries $\{V_m\}$ to be an orthogonal basis for the space of $d \times d$ matrices. Accordingly, the set $\{V_m\}$ will contain d^2 unitaries, labeled by integers from 0 to $d^2 - 1$. For qubits, we will take $\{V_m\}$ to be the four Pauli matrices $\{I, X, Y, Z\}$, labeled as $V_0 = I$, $V_1 = X$, $V_2 = Y$, and $V_3 = Z$.

In terms of the probability distribution $p(m, n)$, the condition (8) amounts to requiring that the marginal probability distributions $p_1(m)$ and $p_2(n)$ be uniform, that is

$$p_1(m) = p_2(n) = \frac{1}{d^2}, \quad \forall m, n \in \{0, ..., d^2 - 1\}. \qquad (10)$$

The probability distributions $p(m, n)$ satisfying Eq. (10) form a convex polytope whose extreme points are probability distributions of the form $p(m, n) = \delta_{m,\sigma(n)}/d^2$, where σ is a permutation of the set $\{0, \ldots, d^2 - 1\}$ [60].

For the identity permutation, satisfying $\sigma(m) = m$ for all values of m, the probability distribution $p(m, n)$ is symmetric, and the interference term (7) is the completely depolarizing channel $\mathcal{G}(\rho) = I/d \ \forall \rho$. Hence, the channel $\mathcal{C}_\omega$ in Eq. (9) is completely depolarizing, and no information can be transmitted through it, no matter what state ω is used. In the following, we will show that, instead, other types of permutations enable a perfect transmission of classical information.

B. Perfect communication through correlated completely depolarizing channels

Here we focus on the case where the message is a qubit ($d = 2$). Let σ be a permutation that swaps two pairs of indices, for example, mapping (0,1,2,3) into (1,0,3,2). In this case, the probability distribution $p(m, n) = \delta_{m,\sigma(n)}/4$ is

symmetric, and the interference term is

$$\mathcal{G}(\rho) = \frac{\rho X e^{i(\phi_1-\phi_0)} + Y\rho Z e^{i(\phi_3-\phi_2)} + \text{h.c.}}{4}, \qquad (11)$$

where h.c. denotes the Hermitian conjugate of the preceding matrices.

Note that $\mathcal{G}(\rho)$ depends only on the differences $\phi_1 - \phi_0$ and $\phi_3 - \phi_2$. We now show that, by suitably choosing the differences $\phi_1 - \phi_0$ and $\phi_3 - \phi_2$, and the state ω, it is possible to achieve a perfect transmission of classical information. When $\phi_1 - \phi_0 = 0$ and $\phi_3 - \phi_2 = \pi/2$, the interference term becomes

$$\mathcal{G}(\rho) = \frac{\{\rho, X\} - \{Z\rho Z, X\}}{4}, \qquad (12)$$

where $\{A, B\} = AB + BA$ denotes the anticommutator of two generic operators A and B. In particular, choosing $\rho = |\pm\rangle\langle\pm|$, with $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$, we obtain

$$\mathcal{G}(|\pm\rangle\langle\pm|) = \pm\frac{I}{2}. \qquad (13)$$

Combining this relation with the depolarizing condition $\mathcal{C}(|\pm\rangle\langle\pm|) = I/2$, and inserting these two relations into into Eq. (5), we obtain

$$\mathcal{C}_\omega(|\pm\rangle\langle\pm|) = \frac{I}{2} \otimes \omega_\pm, \qquad (14)$$

with $\omega_+ := \omega$ and $\omega_- := Z\omega Z$. In other words, the net effect of the superposition of correlated depolarizing channels is to transfer information from the message to the output state of the control.

Putting the control in the state $\omega = |+\rangle\langle+|$, one obtains the orthogonal output states $\omega_\pm = |\pm\rangle\langle\pm|$. Hence, a sender can encode a bit into the states $|\pm\rangle$, and a receiver will be able to decode the bit in principle without error, by measuring the control system in the basis $\{|+\rangle, |-\rangle\}$.

In summary, there exist time-correlated channels that look completely depolarizing when the message is sent at any definite moment of time, and yet allow for a perfect transmission of classical information by sending messages at a coherent superposition of different times.

C. Maximum capacity in the lack of correlations

We now show that correlations in the probability distribution $p(m, n)$ are essential in order to achieve the perfect communication task discussed in the previous subsection. Specifically, we prove that no perfect communication is possible in the lack of correlations, that is, when the probability distribution factorizes as $p(m, n) = p_1(m) p_2(n) = 1/d^4$ [cf. Eq. (10)]. For qubit messages ($d = 2$), we show that, in the lack of correlations,

(1) the classical capacity of the channel $\mathcal{C}_\omega$ is upper bounded by 0.5 bits, meaning that it is impossible to transmit more than 0.5 bits per use of the channel;

(2) the maximum classical capacity of the channel $\mathcal{C}_\omega$ over arbitrary states ω of the control system and over arbitrary (not necessarily random-unitary) realizations of the completely depolarizing channel is equal to 0.16 bits.

The first result follows from an analytical upper bound on the classical capacity, while the second result follows from numerical optimization.

1. Analytical bound on the classical capacity

The derivation of the bound consists of three steps, the details of which are provided in Appendix B.

The first step is to prove that, in the lack of correlations and for message dimension $d = 2$, the channel $\mathcal{C}_\omega$ is entanglement breaking [61], i.e., it transforms all entangled states into separable states. For entanglement-breaking channels, it is known that the classical capacity coincides with the *Holevo capacity* [62]. For a generic quantum channel $\mathcal{E}$, the Holevo capacity is $\chi(\mathcal{E}) = \max_{\{p_x, \rho_x\}} H[\sum_x p_x \mathcal{E}(\rho_x)] - \sum_x p_x \mathcal{E}(\rho_x)$, where the maximum is over all possible ensembles $\{p_x, \rho_x\}$ consisting of a probability distribution $\{p_x\}$ and a set of density matrices $\{\rho_x\}$, and $H(\rho) := -\text{Tr}[\rho \log \rho]$ is the von Neumann entropy of a generic state ρ, log denoting the logarithm in base 2.

The second step is to observe that the state of the control that maximizes the Holevo capacity of the channel $\mathcal{C}_\omega$ is $\omega = |+\rangle\langle+|$. This result holds for arbitrary message dimension $d \geqslant 2$, and, in fact, it holds even in the presence of correlations, as long as the probability distribution $p(m, n)$ is symmetric.

Finally, the third step is to show that, in the lack of correlations and for arbitrary message dimension $d \geqslant 2$, the Holevo capacity of the channel $\mathcal{C}_{|+\rangle\langle+|}$ is upper bounded by $1/d$.

Putting the three steps together, we obtain that, in the lack of correlations and for qubit messages, the classical capacity of the channel $\mathcal{C}_\omega$ is upper bounded by $\frac{1}{2}$ for every possible state ω. Hence, the perfect transmission of 1 bit achieved in Sec. III B is impossible in the lack of correlations.

2. Numerical evaluation of the capacity

The evaluation of the Holevo capacity involves an optimization over all possible input ensembles. For quantum channels with d-dimensional input, the optimization can be restricted to ensembles with up to d^2 linearly independent pure states [63]. In practice, however, the optimization is often hard to carry out even in dimension $d = 2$. To make the optimization feasible, we first show that in our case the optimization can be reduced to an optimization over ensembles that depend only on three real parameters $q, p_0, p_1 \in [0, 1]$. The proof of this result is provided in Appendix C.

Building on the above results, we can numerically evaluate the largest value of the Holevo capacity, and therefore the classical capacity, for all possible qubit channels (i.e., $d = 2$) of the form (4) with $p(m, n) = \frac{1}{16}$. We set the state of the control to $\omega = |+\rangle\langle+|$, which we know to guarantee the maximum Holevo information (cf. Lemma 2 in Appendix B).

The resulting value of the Holevo capacity is a function of the phases $\{\phi_m\}_{m\in\{0,1,2,3\}}$ in Eq. (7). One phase, say ϕ_0, can be set to 0 without loss of generality, as it represents a global phase. In Fig. 4(a), we provide a 3-dimensional plot showing the exact values of the Holevo information, and therefore by the arguments above, the classical capacity, for all possible values of the phases ϕ_1, ϕ_2, and ϕ_3. The maximum over all possible choices of phases is 0.16 bits.
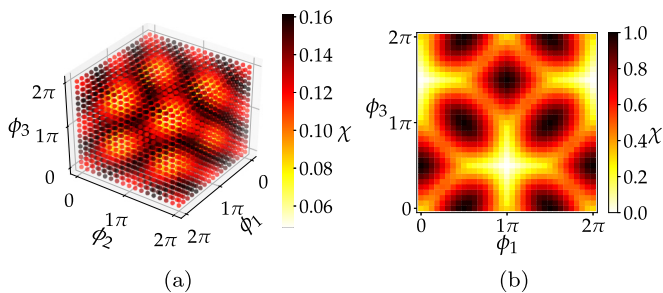
(a)　　　　　　　　(b)

FIG. 4. Performance in the transmission of a single particle through a correlated depolarizing channel. (a) Classical capacity in the lack of correlations. Without loss of generality, $\phi_0 = 0$. The maximum capacity is 0.016 bits. (b) Lower bound to the classical capacity achieved with the correlated probability distribution $p(m,n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. Without loss of generality, we set $\phi_0 = \phi_2 = 0$. The maximum lower bound is 1 bit.

In Appendix C we also show that 0.16 bits is the maximum capacity achievable with arbitrary (not necessarily random unitary) channels that reduce to the depolarizing channel in the one-particle subspace sector. The value 0.16 was previously found to be a lower bound to the classical capacity [31], and our result shows that the lower bound is actually tight: 0.16 is the best classical capacity one can obtain by sending a single particle through a superposition of paths traversing two identical, independent channels that are completely depolarizing in the one-particle subspace.

D. Lower bound to the classical capacity in the presence of correlations

In the correlated case, we do not have a proof that the classical capacity coincides with the Holevo capacity. On top of that, the evaluation of the Holevo capacity generally requires an optimization over all possible ensembles of d^2 linearly independent pure states, which is computationally challenging. Here, we circumvent this problem by computing a lower bound to the Holevo capacity, obtained by restricting the optimization to the set of all *orthogonal* ensembles, that is, input ensembles consisting of two orthogonal qubit states. In general, this lower bound may not be tight [64–66], but it is nevertheless interesting as it quantifies the maximum performance of a natural set of encoding strategies. Since the Holevo capacity is always a lower bound to the classical capacity, the above lower bound is also a lower bound to the classical capacity.

Here, we evaluate the lower bound for the correlated channel with $p(m,n) = \delta_{\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. This particular choice is interesting because as we have seen in Sec. III B, it can reach the maximum capacity of 1 bit. We now inspect how the lower bound depends on the phases.

Since the interference term (11) depends only on the differences $\phi_1 - \phi_0$ and $\phi_3 - \phi_2$, we set $\phi_0 = \phi_2 = 0$ and scan the possible values of ϕ_1 and ϕ_3. For the state of the control system, we choose again $\omega = |+\rangle\langle+|$, as it maximizes the Holevo capacity (cf. Lemma 2 in Appendix B). The lower
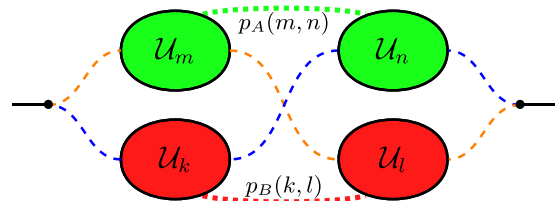

FIG. 5. A single particle is sent through a superposition of two paths (orange and blue dashed lines), each traversing two independent channels (green and red ovals), each of which exhibits time correlations between successive uses. The green and red dotted lines represent the correlations between the two subsequent uses of the same channel.

bound to the Holevo capacity is shown in Fig. 4(b) for all values of ϕ_1 and ϕ_3.

IV. COMMUNICATION THROUGH MULTIPLE TIME-CORRELATED CHANNELS

Time-correlated channels can be used to mimic the use of ordinary quantum channels in a superposition of different causal orders [32,35]. In this section we show that time correlations are a necessary resource for reproducing the benefits of the superposition of orders in quantum communication, and that, in fact, time correlations are an even more powerful resource than the ability to combine channels in a superposition of orders.

A. A network of time-correlated channels

Suppose that two time-correlated channels $\mathcal{R}_A$ and $\mathcal{R}_B$, each of the form (1), are arranged as in Fig. 5, and that a single particle is sent through a superposition of two alternative paths visiting each of the two channels exactly once. When the control system is initialized in the state ω, the overall evolution of the message and the control is described by the *effective channel* $\mathcal{E}_\omega$ defined as

$$\mathcal{E}_\omega(\rho) := \sum_{m,n,k,l} p_A(m,n)p_B(k,l)\, W_{mnkl}(\rho \otimes \omega)W_{mnkl}^\dagger, \quad (15)$$

with

$$
\begin{aligned}
W_{mnkl} := {}& V_l^{(B)}V_m^{(A)}\, e^{i(\phi_k^{(B)}+\phi_n^{(A)})} \otimes |0\rangle\langle 0| \\
& + V_n^{(A)}V_k^{(B)}\, e^{i(\phi_m^{(A)}+\phi_l^{(B)})} \otimes |1\rangle\langle 1|.
\end{aligned}
\quad (16)
$$

Here, $p_A(m,n)$, $p_B(k,l)$, $\{V_m^{(A)}\}$, $\{V_l^{(B)}\}$, $\{\phi_m^{(A)}\}$, and $\{\phi_l^{(B)}\}$ are defined as in Eqs. (1) and (2). The derivation of Eq. (15) is provided in Appendix D.

An interesting special case occurs when the probability distributions $p_A(m,n)$ and $p_B(k,l)$ are perfectly correlated, that is,

$$
\begin{aligned}
p_A(m,n) &= p_{1A}(m)\delta_{mn}, \quad \forall\, m,n \\
p_B(k,l) &= p_{1B}(k)\delta_{kl}, \quad \forall\, k,l
\end{aligned}
\quad (17)
$$

where $p_{1A}(m)$ and $p_{1B}(k)$ are the marginal probability distributions of $p_A(m,n)$ and $p_B(k,l)$, respectively. Under this condition, the network in Fig. 5 reproduces the action of two

random unitary channels in a superposition of two alternative orders [32].

Mathematically, the operation of putting two quantum channels in a superposition of orders is described by the quantum SWITCH [36,37], a higher-order transformation that takes as inputs two generic channels $\mathcal{A}$ and $\mathcal{B}$ (with d-dimensional input and output systems) and produces as output a new quantum channel $\mathcal{S}(\mathcal{A}, \mathcal{B})$ with Kraus operators

$$S_{mk} := A_m B_k \otimes |0\rangle\langle 0| + B_k A_m \otimes |1\rangle\langle 1|, \tag{18}$$

where $\{A_m\}$ ($\{B_k\}$) are Kraus operators of $\mathcal{A}$ ($\mathcal{B}$), and $\{|0\rangle, |1\rangle\}$ is a basis for a control qubit that determines the relative order between $\mathcal{A}$ and $\mathcal{B}$. Notably, the overall channel $\mathcal{S}(\mathcal{A}, \mathcal{B})$ is independent of the choice of Kraus representations for the input channels $\mathcal{A}$ and $\mathcal{B}$.

When the control qubit is put in a fixed state ω, the quantum SWITCH of channels $\mathcal{A}$ and $\mathcal{B}$ yields the effective channel

$$\mathcal{S}_\omega(\rho) := \sum_{m,k} S_{mk}(\rho \otimes \omega)S_{mk}^\dagger, \tag{19}$$

with S_{mk} as in Eq. (18). In particular, here we are interested in the case where the channels $\mathcal{A}$ and $\mathcal{B}$ are random unitary, with Kraus operators $A_m := \sqrt{p_{1A}(m)}\, V_m^{(A)}$ and $B_k := \sqrt{p_{1B}}\, V_k^{(B)}$. With this choice, the channel $\mathcal{S}_\omega$ in Eq. (19) coincides with the channel $\mathcal{E}_\omega$ in Eq. (15) under the condition that the probability distributions $p_A(m, n)$ and $p_B(k, l)$ are perfectly correlated [cf. Eq. (17)].

When the channels $\mathcal{A}$ and $\mathcal{B}$ are completely depolarizing, Ref. [45] showed that the channel $\mathcal{S}_\omega$ resulting from the quantum SWITCH can transmit 0.049 bits of classical information, provided that the control is initialized in the state $\omega = |+\rangle\langle +|$. Later, the value 0.049 was proven to be exactly equal to the classical capacity [50]. Since the channels $\mathcal{E}_\omega$ and $\mathcal{S}_\omega$ coincide, we conclude that the time-correlated network in Fig. 5 can achieve a capacity of 0.049 bits.

In the following, we provide two results:

(1) We show that time correlations are strictly necessary in order to achieve the quantum SWITCH capacity of 0.049 bits. Specifically, we show numerically that the maximum classical capacity in the uncorrelated case is 0.018 bits for random-unitary realizations of the completely depolarizing channel, and 0.024 bits for arbitrary realisations. This result shows that, when the quantum SWITCH is reproduced by the network in Fig. 5, the origin of the communication enhancement is not just the interference of paths, but rather the combined effect of the interference of paths *and* of the time correlations.

(2) We show that there exist time correlations that achieve a classical capacity of at least 0.31 bits. This result shows that the access to time correlations is generally a stronger resource than the ability to combine ordinary channels in a superposition of orders.

B. Maximum capacity in the lack of correlations

Here we evaluate the maximum amount of classical information that can be transmitted through the network in Fig. 5 when the channels are completely depolarizing and no correlation is present, that is, when $p_A(m, n) = p_B(k, l) = 1/16 \; \forall \; m, n, k, l \in \{0, 1, 2, 3\}$.
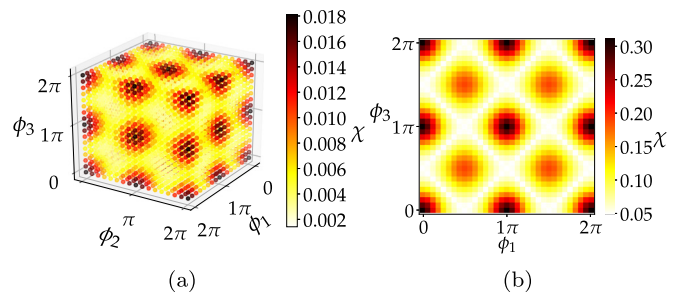

FIG. 6. Performance in the transmission of a single particle through a network of correlated depolarizing channels, arranged as in Fig. 5. (a) Classical capacity in the lack of correlations. Without loss of generality, $\phi_0 = 0$. The maximum capacity is 0.018 bits. (b) Lower bound to the classical capacity achieved with maximal correlations corresponding to the probability distributions $p_A(m, n) = p_B(m, n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. Without loss of generality, $\phi_0 = \phi_2 = 0$. The maximum lower bound is 0.31 bits.

The evaluation of the maximum capacity follows the same steps as in Sec. III C. The main observations are as follows:

(1) In the lack of correlations, the channel $\mathcal{E}_\omega$ in Eq. (15) is entanglement breaking, and therefore its classical capacity coincides with the Holevo capacity.

(2) The control state ω that maximizes the Holevo capacity of the channel $\mathcal{E}_\omega$ is $\omega = |+\rangle\langle +|$.

(3) Without loss of generality, the maximization of the Holevo information can be reduced to ensembles that depend only on three real parameters q, p_0, and p_1 in [0,1].

The derivation of these results is provided in Appendix E.

Building on the above observations, we evaluate the capacity of the channel $\mathcal{E}_\omega$ in Eq. (15) by scanning all possible values of the phases $\{\phi_m\}_{m=0}^3$. The result is the plot shown in Fig. 6(a). The largest classical capacity over all random unitary realizations is 0.018 bits, which is strictly smaller than the value 0.049 bits achieved by the superposition of orders.

Furthermore, we also extend the optimization from random unitary realizations to arbitrary realizations of the completely depolarizing channel. For this broader class of realizations, we numerically obtain that the maximum capacity is 0.024 bits.

Summarizing, the best classical capacity one can obtain by sending a single particle through the network in Fig. 5, in the lack of correlations between the two paths, is 0.018 bits, and the capacity can be increased to 0.024 bits by replacing the random unitary channels with more general realizations of the completely depolarizing channel.

Note that both values 0.018 and 0.024 are below the 0.049 bits of classical capacity achieved by the quantum SWITCH. This result shows that, when the quantum SWITCH is reproduced by the correlated network in Fig. 5, it offers a communication advantage over all communication protocols where a single particle travels in a superposition of two paths on which it experiences uncorrelated noisy processes. Hence, we conclude that, in this scenario, the origin of the communication advantages of the quantum SWITCH is not merely the superposition of paths, but rather the nontrivial interplay between the superposition of paths and the time correlations in the noise.

Our results also imply a *caveat* about terminology. The quantum SWITCH of two channels $\mathcal{A}$ and $\mathcal{B}$ is sometimes described informally as a *"superposition of channels $\mathcal{AB}$ and $\mathcal{BA}$."* While this expression may be formally correct (at least according to a broad notion of superposition [32]), it can be misleading if taken at face value because it does not mention explicitly the requirement of correlations between the channels $\mathcal{A}$ and $\mathcal{B}$ in the two branches of the superposition.

C. Time correlations surpassing the quantum SWITCH capacity

We now show that the classical capacity of 0.049 bits, achieved by the quantum SWITCH, can be surpassed using more general time correlations. We prove this result explicitly, by exhibiting a pair of time-correlated channels that achieve a capacity at least 0.31 bits.

Our choice of channels corresponds to $p_A(m, n) = p_B(m, n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. This choice is motivated by the fact that the permutation σ guarantees the maximum communication capacity in the case where a single time-correlated channel is used (cf. Sec. III B).

With the above choice, the effective channel describing the transmission of the message is

$$\mathcal{E}_\omega(\rho) = \frac{\frac{I}{2} + \mathcal{K}(\rho)}{2} \otimes \omega + \frac{\frac{I}{2} - \mathcal{K}(\rho)}{2} \otimes Z\omega Z, \quad (20)$$

with

$$\mathcal{K}(\rho) := \frac{1}{8}\{[\cos 2(\phi_1 - \phi_0) + \cos 2(\phi_3 - \phi_2)]\rho + 2X\rho X\}. \quad (21)$$

The derivation of this formula is provided in Appendix D. Note that the channel $\mathcal{E}_\omega$ depends only on the phase differences $\phi_1 - \phi_0$ and $\phi_3 - \phi_2$, via Eq. (21).

We now provide a lower bound to the classical capacity of the channel $\mathcal{E}_\omega$. As we did earlier in the paper, we lower bound the classical capacity by the Holevo capacity, and, in turn, we lower bound the Holevo capacity by restricting the maximization to orthogonal input ensembles. For the state of the control qubit, we pick $\omega = |+\rangle\langle+|$, which is the choice that maximizes the Holevo capacity (cf. Lemma 2 in Appendix B).

The lower bound to the classical capacity is shown in Fig. 6(b) for all possible values of the phase differences $\phi_1 - \phi_0$ and $\phi_3 - \phi_2$. The highest lower bound over all combinations of phases $\{\phi_m\}_{m=0}^3$ is given by 0.31 bits. This value is larger than the classical capacity of 0.049 bits achieved by the quantum SWITCH, corresponding to perfect correlations $p_A(m, n) = p_B(m, n) = \delta_{m,n}/4$. This result implies that not only can time correlations reproduce the superposition of causal orders, but they can also surpass its advantages.

V. NOISE ON THE CONTROL DEGREE OF FREEDOM

So far, we have assumed that the message-carrying degree of freedom of the particle undergoes noise during transmission, while the control degree of freedom is noiseless. However, in practical scenarios, this will only be an approximation to the actual physics. We now briefly discuss the
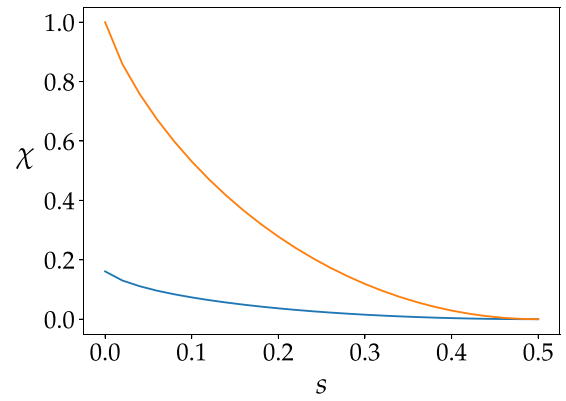

FIG. 7. Blue: Maximum classical capacity in the absence of correlations, as a function of the dephasing parameter s. The maximum is computed over all realizations of the completely depolarizing channel, and is achieved by the random unitary realization with the choice of phases $\{\phi_m\}$ that give the maximum capacity of 0.16 bits when $s = 0$. Orange: Lower bound to the maximal classical capacity in the presence of correlations, as a function of the dephasing parameter s. The lower bound is computed by considering the correlated probability distribution $p(m, n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3, and $\phi_0 = \phi_1 = \phi_2 = 0$, $\phi_3 = \pi/2$.

effect of noise on the control system, focusing in particular on dephasing noise, of the form

$$\mathcal{P}(\omega) = sZ\omega Z + (1 - s)\omega, \quad (22)$$

where $s \in [0, \frac{1}{2}]$ is a probability and ω is the initial state of the control. For a more detailed investigation into the effects of noise on the control system, we refer the reader to a recent related work [67].

For simplicity, here we focus on the communication scenario involving a single transmission line, as in Fig. 1. In this setting, the evolution experienced by a single particle is described by the channel

$$\mathcal{C}'_\omega := (\mathcal{I}_M \otimes \mathcal{P})\mathcal{C}_\omega, \quad (23)$$

obtained by dephasing the control system at the output of the channel $\mathcal{C}_\omega$ in Eq. (5). By inserting the expression (5) into the above equation, it is immediate to see that the effect of dephasing is to dampen the interference term $\mathcal{G}$ in the effective channel (5): specifically, the interference term changes from $\mathcal{G}$ to $(1 - 2s)\mathcal{G}$.

In the case of completely depolarizing channels on the message degree of freedom, the presence of a nonzero interference term means that, as long as the dephasing of the control is not complete ($s \neq \frac{1}{2}$), the superposition of evolutions can still allow for a nonzero amount of classical information to be transmitted, thereby offering an advantage over the transmission at a definite moment of time.

Figure 7 shows the behavior of the classical capacity as a function of the dephasing parameter s. The figure shows that correlations between two uses of the channel offer an enhancement of the classical capacity. To make this point, we first evaluate numerically the maximum capacity achievable in the lack of correlations, with arbitrary realizations of the completely depolarizing channel (blue curve). Notably,

the capacity for every fixed value of s is achieved by the same realization of the completely depolarizing channel that achieves the maximum capacity in the ideal $s = 0$ case. We then show that a higher capacity can be achieved with the correlated channel described in Sec. III B. To this purpose, we numerically evaluate a lower bound to the Holevo capacity (and therefore the classical capacity), obtained by restricting the maximization to orthogonal input ensembles (orange curve). Note that both the blue and orange curves are above 0 for every nonmaximal amount of dephasing ($s \neq \frac{1}{2}$), meaning that the single-particle transmission at a coherent superposition of times offers an advantage over the transmission at a definite time.

VI. CONCLUSIONS

We have shown that a single quantum particle can sense the correlations between noisy processes at different moments of time. By sending the particle at a superposition of different times, one can take advantage of these correlations and boost the communication rate to values that would be impossible if the moment of transmission were a classical, well-defined variable.

An important avenue for future research is the experimental realization of our protocols, as well as the experimental exploration of their noise robustness to timing errors and decoherence between the two different modes used to create the superposition. On the theoretical side, it is interesting to apply our framework for single-particle communication to more complex scenarios, e.g., involving the transmission of a single particle at more than two times, or even in continuous time. It is also interesting to analyze other communication tasks, such as the two-way communication proposed in Ref. [68]. Moreover, the extension from single-particle communication to other communication protocols with a finite number of particles is a natural next step of this research.

At the foundational level, time-correlated channels provide an insight into the resources used by the existing experiments on the superposition of causal order. We analyzed a basic setup that reproduces the overall result of the quantum SWITCH by sending a single particle in a superposition of paths through time-correlated channels. In this setup, we showed that time correlations are a necessary resource to reproduce the communication advantages of the quantum SWITCH. Moreover, we observed that, with more elaborate patterns of correlations, one can achieve an even greater enhancement than the one found for the superposition of orders. This result establishes time-correlated channels as an appealing resource, which can be used as a testbed for foundational results on causal order, and, at the same time, as a building block for new communication protocols.

ACKNOWLEDGMENTS

APPENDIX A: TRANSMISSION OF A SINGLE PARTICLE THROUGH A SUPERPOSITION OF MULTIPLE PORTS

Here we provide a mathematical framework for describing the transmission of a single particle at a superposition of different times, and, more generally, for describing the transmission of the particle on a superposition of different trajectories, each passing through one of the ports of a multiport quantum device.

1. Multiport quantum devices and their vacuum extensions

A transmission line with a single input port is described by a quantum channel, that is, a completely positive trace-preserving map transforming density matrices on the particle's Hilbert space. In the following we will denote by $\mathsf{Chan}(S \rightarrow S')$ the set of quantum channels with input system S and (possibly different) output system S'. When $S = S'$ we will use the shorthand $\mathsf{Chan}(S)$. The action of a quantum channel $\mathcal{A}$ on a density matrix ρ can be conveniently written in the Kraus representation $\mathcal{A}(\rho) = \sum_i A_i \rho A_i^\dagger$, where $\{A_i\}$ is a (nonunique) set of operators, satisfying $\sum_i A_i^\dagger A_i = I$.

A transmission line with k input and output ports is described by a k-partite quantum channel $\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes \cdots \otimes S^{(k)} \rightarrow S'^{(1)} \otimes \cdots \otimes S'^{(k)})$ with k input-output pairs $(S^{(i)}, S'^{(i)})_{i=1}^k$.

A transmission line that can be used k times in succession is described by *k-step quantum channel* [6] (also known as a quantum k comb [51,52]). A k-step quantum channel is a special type of k-partite channel $\mathcal{B}$ with the additional property that no signal propagates from an input $S^{(i)}$ to any group of outputs $S'^{(j)}$ with $j < i$ [51]. We will denote the set of k-step quantum channels as $\mathsf{Chan}(S^{(1)} \rightarrow S'^{(1)}, \ldots, S^{(k)} \rightarrow S'^{(k)})$, or simply $\mathsf{Chan}(S^{(1)}, \ldots, S^{(k)})$ when the input and output of each pair coincide. For $k = 2$, an example of two-step quantum channel is illustrated in Fig. 8.

The possibility that no particle is sent through a port of a device can be described using the notion of *vacuum*
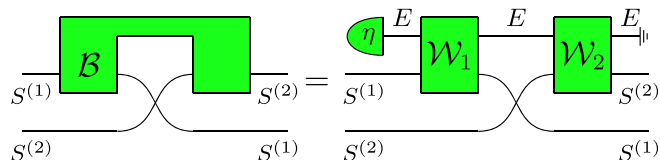
FIG. 8. The left-hand side depicts a two-step correlated quantum channel $\mathcal{B}$ taking two input states on systems $S^{(1)}$ and $S^{(2)}$, in succession. The right-hand side shows the physical implementation of the two-step channel via two unitary channels $\mathcal{W}_1$ and $\mathcal{W}_2$ [51,52] where the memory between the two uses of the channel is realized by an environment E, which is inaccessible to the communicating parties.

extension [32]. Consider first a single-port device, described by an ordinary quantum channel $\mathcal{A} \in \mathsf{Chan}(S)$. When no particle is sent through the device, we describe the input as the vacuum state $|\text{vac}\rangle$, that is, a state in a *vacuum sector* Vac [32,33,69,70], which is orthogonal to the one-particle sector S. Overall, the device acts on an extended system $\widetilde{S} := S \oplus \text{Vac}$, which is associated with the Hilbert space given by $\mathcal{H}_S \oplus \mathcal{H}_{\text{Vac}}$, where $\mathcal{H}_{\text{Vac}}$ is the vacuum Hilbert space, here assumed to be one dimensional.

Given a quantum channel $\mathcal{A}$, a vacuum extension $\widetilde{\mathcal{A}}$ of $\mathcal{A}$ is any channel which acts as $\mathcal{A}$ (respectively, $\mathcal{I}_{\text{Vac}}$) when the input is a state in sector S (respectively, Vac). The Kraus operators of $\widetilde{\mathcal{A}}$ are $\widetilde{A}_i = A_i \oplus \alpha_i |\text{vac}\rangle\langle\text{vac}|$, where $\{A_i\}_{i=0}^{r-1}$ is a Kraus representation of $\mathcal{A}$, and $\{\alpha_i\}_{i=0}^{r-1}$ are *vacuum amplitudes* satisfying $\sum_{i=0}^{r-1} |\alpha_i|^2 = 1$.

A given channel has infinitely many possible vacuum extensions. In an actual communication scenario, the vacuum extension can be determined by probing the action of the channel on superpositions of the vacuum and one-particle states. Physically, the choice of vacuum extension is determined by the Hamiltonian of the field describing the vacuum and the one-particle sector.

The notion of vacuum extension can be easily extended to the case of k-partite channels, which include k-step channels as a special case. For simplicity, we focus on the $k = 2$ case, but the extension to $k \geqslant 2$ is straightforward.

Consider a transmission line described by a bipartite channel $\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes S^{(2)})$. A vacuum extension of the channel $\mathcal{B}$ is another bipartite channel $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)})$, acting on the extended systems $\widetilde{S}^{(1)} := S^{(1)} \oplus \text{Vac}^{(1)}$ and $\widetilde{S}^{(2)} := S^{(2)} \oplus \text{Vac}^{(2)}$. In general, the systems $S^{(1)}, S^{(2)}$ can represent the systems accessible at the same location at two consecutive moments of time, or it can represent the systems accessible at different locations at the same time (as considered in Refs. [31,32]) or, more generally, they can represent any pair of independently addressable systems, representing the input and output ports of our multiport device.

2. A single particle traveling through multiple ports

In order to be able to send the same quantum particle to either of the ports of the device, we require the isomorphism $S^{(1)} \cong S^{(2)} \cong M$, where M is the message-carrying degree of freedom of the particle. In this case, the tensor product $\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)}$ contains a no-particle sector $\text{Vac}^{(1)} \otimes \text{Vac}^{(2)}$,
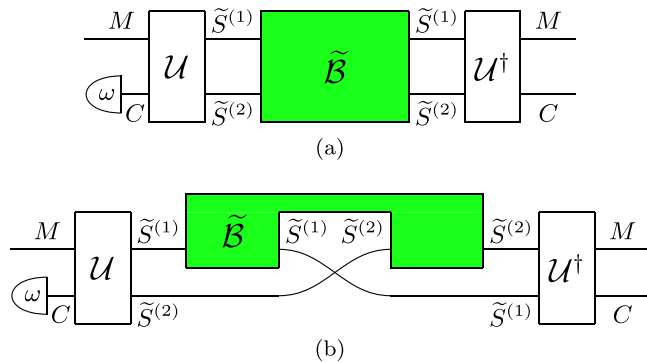

FIG. 9. (a) Transmission of a single particle through a bipartite quantum channel $\widetilde{\mathcal{B}}$ (green). (b) Transmission of a single particle through a two-step quantum channel $\widetilde{\mathcal{B}}$ (green). In both cases, the particle is represented by a composite system $M \otimes C$, where M represents the degrees of freedom used as the message, and C represents the degrees of freedom used as the control. The isomorphism $\mathcal{U}$ converts the composite system $M \otimes C$ into the one-particle sector $(S^{(1)} \otimes \text{Vac}) \oplus (\text{Vac} \otimes S^{(2)})$ of $\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)}$. The inverse map $\mathcal{U}^{\dagger}$ converts the output state back into $M \otimes C$. For the applications in this paper, we take the input of the control system C to be fixed in the state ω while the message system M is accessible to the sender.

a one-particle sector $(S^{(1)} \otimes \text{Vac}^{(2)}) \oplus (\text{Vac}^{(1)} \otimes S^{(2)})$, and a two-particle sector $S^{(1)} \otimes S^{(2)}$. The one-particle sector is isomorphic to $M \otimes C$, where C is a qubit system, representing the degree of freedom of the particle that controls its time of transmission. When the control is in state $|0\rangle$, the message is sent through the first application of the channel and the vacuum is sent in the second application; vice versa for the control in state $|1\rangle$.

We now define the situation in which a single particle is sent at a superposition of two different ports. We call the process experienced by the particle the *superposition channel* $\mathcal{S}(\widetilde{\mathcal{B}})$, and define it as the restriction of $\widetilde{\mathcal{B}}$ to the one-particle sector, regarded as isomorphic to the composite system "message + control." Explicitly, the action of the superposition channel is defined as

$$\mathcal{S}(\widetilde{\mathcal{B}}) := \mathcal{U}^{\dagger} \circ \widetilde{\mathcal{B}} \circ \mathcal{U}, \tag{A1}$$

where $\mathcal{U}(\cdot) := U(\cdot)U^{\dagger}$ is the isomorphism between $M \otimes C$ and the one-particle sector $(S^{(1)} \otimes \text{Vac}) \oplus (\text{Vac} \otimes S^{(2)})$, with

$$U(|\psi\rangle_M \otimes |0\rangle_C) := |\psi\rangle_{\widetilde{S}^{(1)}} \otimes |\text{vac}\rangle_{\widetilde{S}^{(2)}},$$
$$U(|\psi\rangle_M \otimes |1\rangle_C) := |\text{vac}\rangle_{\widetilde{S}^{(1)}} \otimes |\psi\rangle_{\widetilde{S}^{(2)}}. \tag{A2}$$

Mathematically, the transformation $\mathcal{S} : \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)}) \to \mathsf{Chan}(M \otimes C)$ is a quantum supermap, that is, a transformation from quantum channels to quantum channels satisfying appropriate consistency requirements [37,52,71]. An illustration of the supermap $\mathcal{S}$ is provided in Fig. 9(a).

Note that definition (A1) can be applied in particular to k-step quantum channels, which are a special case of k-partite channels. The illustration of the supermap $\mathcal{S}$ in this special case is provided in Fig. 9(b).

The same definition can be adopted for the transmission of a single particle through a k-partite multiport device. In this case, the device is represented by a k-partite quantum channel

$\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes \cdots \otimes S^{(k)})$, with $S^{(1)} \cong S^{(2)} \cong \cdots \cong S^{(k)}$, and with vacuum extension $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \cdots \otimes \widetilde{S}^{(k)})$. The superposition channel is then defined as the restriction of $\widetilde{\mathcal{B}}$ to the one-particle sector

$$
\bigoplus_{j=1}^{k} \mathrm{Vac}^{(1)} \otimes \cdots \otimes \mathrm{Vac}^{(j-1)} \otimes S^{(j)} \otimes \mathrm{Vac}^{(j+1)} \otimes \cdots \otimes \mathrm{Vac}^{(k)}
$$
$$
\cong M \otimes C, \tag{A3}
$$

where C is now a k-dimensional control system.

3. Derivation of Eq. (4) in the main text

We now specialize to the case of correlated channels of the random unitary form

$$
\mathcal{R} = \sum_{m,n} p(m,n) \, \mathcal{V}_m \otimes \mathcal{V}_n \in \mathsf{Chan}(S^{(1)}, S^{(2)}), \tag{A4}
$$

where $\mathcal{V}_m(\cdot) := V_m(\cdot)V_m^{\dagger}$ is a unitary channel, $\{V_m\}$ is a set of unitary gates, and $p(m,n)$ is a joint probability distribution. The vacuum extension of each unitary V_m is taken to be another unitary U_m, which we write as

$$
\widetilde{V}_m := U_m = V_m \oplus e^{i\phi_m} |\mathrm{vac}\rangle\langle\mathrm{vac}|, \tag{A5}
$$

where the vacuum amplitude is given by a complex phase, representing the coherent action of each possible noisy process on the one-particle and vacuum sectors. This leads to the vacuum extension

$$
\widetilde{\mathcal{R}} = \sum_{m,n} p(m,n) \, \widetilde{\mathcal{V}}_m, \quad \widetilde{\mathcal{V}}_n \in \mathsf{Chan}(\widetilde{S}^{(1)}, \widetilde{S}^{(2)}), \tag{A6}
$$

with $\widetilde{\mathcal{V}}_m(\cdot) := \widetilde{V}_m(\cdot)\widetilde{V}_m^{\dagger}$, which is equivalent to Eq. (1) in the main text, with $U_m = \widetilde{V}_m$.

The use of the channel $\mathcal{R}$, specified by the vacuum extension $\widetilde{\mathcal{R}}$, at a superposition of times is given by

$$
\mathcal{S}(\widetilde{\mathcal{R}}) = \sum_{m,n=0}^{r-1} p(m,n) \, \mathcal{U}^{\dagger} \circ (\widetilde{\mathcal{V}}_m \otimes \widetilde{\mathcal{V}}_n) \circ \mathcal{U}. \tag{A7}
$$

Explicitly, we have the expression

$$
\mathcal{S}(\widetilde{\mathcal{R}})(\rho \otimes \omega) = \sum_{m,n} C_{mn} (\rho \otimes \omega) C_{mn}^{\dagger}, \tag{A8}
$$

where ρ (respectively, ω) is an arbitrary state of the message (respectively, control), and

$$
C_{mn} := \sqrt{p(m,n)} e^{i\phi_n} V_m \otimes |0\rangle\langle 0|
$$
$$
+ \sqrt{p(m,n)} V_n e^{i\phi_m} \otimes |1\rangle\langle 1|, \tag{A9}
$$

$e^{i\phi_m}$ being the vacuum amplitude in Eq. (A5). Equation (A8) coincides with Eq. (4) in the main text, with $\mathcal{C} := \mathcal{S}(\widetilde{\mathcal{R}})$ and $W_{mn} := C_{mn}/\sqrt{p(m,n)}$.

4. Derivation of Eqs. (5)–(7) in the main text

It is useful to consider the case where the probability distribution $p(m,n)$ is symmetric, that is, $p(m,n) = p(n,m)$ for every m and n. In this case, the superposition channel has the

simple expression

$$
\mathcal{S}(\widetilde{\mathcal{R}}) = \frac{\mathcal{R}_1 + \mathcal{G}}{2} \otimes \mathcal{I} + \frac{\mathcal{R}_1 - \mathcal{G}}{2} \otimes \mathcal{Z}, \tag{A10}
$$

where $\mathcal{Z}$ is the unitary channel associated to the Pauli matrix Z, $\mathcal{R}_1$ is the reduced channel defined by

$$
\mathcal{R}_1(\rho) := \sum_m p_1(m) V_m \rho V_m^{\dagger}, \quad p_1(m) := \sum_n p(m,n), \tag{A11}
$$

and $\mathcal{G}$ is the linear map defined by

$$
\mathcal{G}(\rho) := \sum_{m,n} p(m,n) \, e^{i(\phi_n - \phi_m)} V_m \rho V_n^{\dagger}. \tag{A12}
$$

APPENDIX B: ANALYTICAL BOUND ON THE CLASSICAL CAPACITY IN THE LACK OF CORRELATIONS

This section refers to the scenario where the message is transmitted at a superposition of two possible times, experiencing independent noisy processes that are completely depolarizing in the one-particle subspace. This section makes use of the notation introduced in Appendix A.

1. Proof that the superposition of uncorrelated completely depolarizing channels is entanglement breaking

Let $\mathcal{A}(\cdot) = \sum_{m=0}^{r-1} A_m(\cdot)A_m^{\dagger} \in \mathsf{Chan}(S)$ be a generic quantum channel, and let $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{S})$ be a vacuum extension of $\mathcal{A}$. Using Eq. (A1), we obtain

$$
\mathcal{S}(\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{A}}) = \frac{\mathcal{A}(\rho) + F\rho F^{\dagger}}{2} \otimes \mathcal{I} + \frac{\mathcal{A}(\rho) - F\rho F^{\dagger}}{2} \otimes \mathcal{Z}, \tag{B1}
$$

where $\mathcal{I}$ (respectively, $\mathcal{Z}$) is the identity channel (respectively, Pauli channel corresponding to the Pauli matrix Z), and

$$
F := \sum_m \overline{\alpha}_m A_m \tag{B2}
$$

is the *vacuum interference operator* defined in Ref. [32].

Now, let $\mathcal{A}$ be the completely depolarizing channel $\mathcal{D} : \rho \mapsto I/d$, with vacuum extension $\widetilde{\mathcal{D}}$. For a fixed state ω of the control system, consider the effective channel defined by

$$
\mathcal{S}(\widetilde{\mathcal{D}} \otimes \widetilde{\mathcal{D}})(\rho \otimes \omega) = \frac{I/d + F\rho F^{\dagger}}{2} \otimes \mathcal{I} + \frac{I/d - F\rho F^{\dagger}}{2} \otimes \mathcal{Z}
$$
$$
=: \mathcal{C}_{\omega, F}(\rho). \tag{B3}
$$

For $d = 2$, we have the following result:

Proposition 1. The channel $\mathcal{C}_{\omega, F}$ in Eq. (B3) is entanglement breaking for $d = 2$.

The proof uses the following lemma:

Lemma 1. Let $\mathcal{D}$ be a completely depolarizing channel with vacuum extension $\widetilde{\mathcal{D}}$ and vacuum interference operator F. Then, the operator norm of F satisfies the inequality $||F||_{\infty} \leqslant \frac{1}{\sqrt{d}}$.

Proof. Let the Kraus operators and vacuum amplitudes of $\mathcal{D}$ be given by $\{A_i\}$, $\{\alpha_i\}$, respectively. By definition,

$$
||F||_{\infty} = \max_{\{|v\rangle : |||v\rangle||=1\}} \max_{\{|w\rangle : |||w\rangle||=1\}} \langle v|F|w\rangle \tag{B4}
$$

and

$$
|\langle v|F|w\rangle| = \left| \sum_i \overline{\alpha_i} \langle v|A_i|w\rangle \right|
$$

$$
\leqslant \sqrt{ \left(\sum_i |\alpha_i|^2 \right) \left(\sum_j \langle v|A_j|w\rangle \langle w|A_j^\dagger|v\rangle \right) }
$$

$$
= \sqrt{\langle v|\mathcal{D}(|w\rangle\langle w|)|v\rangle}. \tag{B5}
$$

If $\mathcal{D}$ is the completely depolarizing channel, then $\mathcal{D}(|w\rangle\langle w|) = I/d$ and therefore the bound becomes $|\langle v|F|w\rangle| \leqslant \sqrt{1/d}$ which implies $||F||_\infty \leqslant \sqrt{1/d}$. ∎

We are now ready to provide the proof of Proposition 1.

Proof of Proposition 1. To prove that a channel is entanglement breaking, it is sufficient show that it transforms a maximally entangled state into a separable state [61]. Let $|\Phi^+\rangle = \sum_{k=0}^{d-1} |k\rangle \otimes |k\rangle/\sqrt{d}$ be the canonical maximally entangled state. When the channel $\mathcal{C}_{\omega,F}$ is applied, the output state is

$$
(\mathcal{C}_{\omega,F} \otimes \mathcal{I})(|\Phi^+\rangle\langle\Phi^+|) = \left(\frac{I \otimes I}{d^2} + G_F \right) \otimes \frac{\omega}{2}
$$

$$
+ \left(\frac{I \otimes I}{d^2} - G_F \right) \otimes \frac{Z\omega Z}{2}, \tag{B6}
$$

with $G_F := (F \otimes I)(|\Phi^+\rangle\langle\Phi^+|)(F \otimes I)^\dagger$.

We now show that the operators $\frac{I \otimes I}{d^2} \pm G_F$ are proportional to states with positive partial transpose. To this purpose, note that the partial transpose of G_F on the second space is

$$
G_F^{\tau_2} = (F \otimes I) \frac{\mathtt{SWAP}}{d} (F \otimes I)^\dagger. \tag{B7}
$$

Hence, for every unit vector $|\Psi\rangle$ we have the bound

$$
\langle\Psi|G_F^{\tau_2}|\Psi\rangle \leqslant \frac{\langle\Psi|(FF^\dagger \otimes I)|\Psi\rangle}{d} \leqslant \frac{\|FF^\dagger\|_\infty}{d}
$$

$$
= \frac{\|F\|_\infty^2}{d} \leqslant \frac{1}{d^2}, \tag{B8}
$$

where the first inequality follows from Schwarz' inequality, and the last inequality follows from Lemma 1.

Using Eq. (B8), we obtain the relation

$$
\langle\Psi| \left(\frac{I \otimes I}{d^2} \pm G_F \right)^{\tau_2} |\Psi\rangle \geqslant \frac{1}{d^2} - \langle\Psi|G_F^{\tau_2}|\Psi\rangle \geqslant 0. \tag{B9}
$$

Since $|\Psi\rangle$ is an arbitrary vector, we conclude that the operator $(\frac{I \otimes I}{d^2} \pm G_F)^{\tau_2}$ has positive partial transpose. For $d = 2$, the Peres-Horodecki criterion [72,73], guarantees that $\frac{I \otimes I}{4} \pm G_F$ is proportional to a separable state. Hence, the whole output state (B6) is separable. ∎

2. Optimal control state for maximizing the Holevo capacity

Proposition 1 implies that the classical capacity of the channel $\mathcal{C}_{\omega,F}$ is equal to its Holevo capacity (see [62]). Here we show that the Holevo capacity is maximized by the state $\omega = |+\rangle\langle+|$. In fact, we prove a more general result:

Lemma 2. Let $\mathcal{C}_\omega$ be an arbitrary channel of the form

$$
\mathcal{C}_\omega(\rho) := \mathcal{L}_+(\rho) \otimes \omega + \mathcal{L}_-(\rho) \otimes Z\omega Z, \tag{B10}
$$

where $\mathcal{L}_\pm$ are arbitrary linear maps. Then, for every density matrix ω, the Holevo capacity satisfies the bound $\chi(\mathcal{C}_\omega) \leqslant \chi(\mathcal{C}_{|+\rangle\langle+|})$.

Proof. The Holevo capacity is known to be monotonically decreasing under the composition of quantum channels, namely, $\chi(\mathcal{E}) \geqslant \chi(\mathcal{F} \circ \mathcal{E})$ for every pair of channels $\mathcal{E}$ and $\mathcal{F}$. For every channel $\mathcal{C}_\omega$ of the form (B10), we have the relation

$$
\mathcal{C}_\omega = (\mathcal{I}_M \otimes \mathcal{P}_\omega) \circ \mathcal{C}_{|+\rangle\langle+|}, \tag{B11}
$$

where $\mathcal{P}_\omega$ is the quantum channel defined by

$$
\mathcal{P}_\omega(\gamma) := \langle+|\gamma|+\rangle \, \omega + \langle-|\gamma|-\rangle \, Z\omega Z \tag{B12}
$$

for an arbitrary state γ. Hence, we have $\chi(\mathcal{C}_\omega) = \chi[(\mathcal{I}_M \otimes \mathcal{P}_\omega) \circ \mathcal{C}_{|+\rangle\langle+|}] \leqslant \chi(\mathcal{C}_{|+\rangle\langle+|})$. ∎

Lemma 2 holds in particular for
(1) the channel $\mathcal{C}_{\omega,F}$ defined in Eq. (B3),
(2) the channel $\mathcal{C}_\omega$ defined in Eq. (9) of the main text,
(3) the channel $\mathcal{E}_{\omega,F}$ defined in Eq. (20) of the main text.

3. Bound on the Holevo capacity

Proposition 2. The Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ defined in Eq. (B3) is upper bounded as

$$
\chi(\mathcal{C}_{\omega,F}) \leqslant \frac{\log(2d)}{d} + \frac{\frac{1}{d} + \|F\|_\infty^2}{2} \log \frac{\frac{1}{d} + \|F\|_\infty^2}{2}
$$

$$
+ \frac{\frac{1}{d} - \|F\|_\infty^2}{2} \log \frac{\frac{1}{d} - \|F\|_\infty^2}{2}, \tag{B13}
$$

where F is the vacuum interference operator defined in Eq. (B2).

Proof. For a fixed vacuum extension, and therefore for a fixed vacuum interference operator F, the Holevo capacity of the channel $\mathcal{C}_\omega$ is upper bounded by the Holevo capacity of the channel $\mathcal{C}_{|+\rangle\langle+|,F}$ (Lemma 2). Hence, it is enough to prove the bound for the channel $\mathcal{C}_{|+\rangle\langle+|}$.

Note that the output of channel $\mathcal{C}_{|+\rangle\langle+|,F}$ has dimension $2d$. For a generic channel $\mathcal{E}$ with $(2d)$-dimensional output, the Holevo capacity is upper bounded as [74]

$$
\chi(\mathcal{E}) \leqslant \log(2d) - \min_\rho H[\mathcal{E}(\rho)], \tag{B14}
$$

where $H(\rho) := -\mathrm{Tr}[\rho \log \rho]$ is the von Neumann entropy, and the minimization can be restricted without loss of generality to pure states.

We now upper bound the right-hand side of Eq. (B14) for $\mathcal{E} = \mathcal{C}_{|+\rangle\langle+|,F}$. The action of the channel $\mathcal{C}_{|+\rangle\langle+|,F}$ on a generic input state ρ is

$$
\mathcal{C}_{|+\rangle\langle+|,F}(\rho) = \frac{\frac{I}{d} + F\rho F^\dagger}{2} \otimes |+\rangle\langle+| + \frac{\frac{I}{d} - F\rho F^\dagger}{2} \otimes |-\rangle\langle-|, \tag{B15}
$$

as one can deduce from Eqs. (B3) and (B1).

In the case of a pure state $\rho = |\psi\rangle\langle\psi|$, we write $F|\psi\rangle = k|\varphi\rangle$, where $|\varphi\rangle$ is a unit vector and k is a normalization

constant. With this notation, we obtain

$$
\mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|) = \frac{\left(\frac{1}{d}+k^2\right)|\varphi\rangle\langle\varphi| + \frac{1}{d}P_\perp}{2} \otimes |+\rangle\langle+|
$$
$$
+ \frac{\left(\frac{1}{d}-k^2\right)|\varphi\rangle\langle\varphi| + \frac{1}{d}P_\perp}{2} \otimes |-\rangle\langle-|,
$$
(B16)

with $P_\perp := I - |\varphi\rangle\langle\varphi|$. The von Neumann entropy of this state is

$$
H[\mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|)]
$$
$$
= -\frac{\frac{1}{d}+k^2}{2}\log\frac{\frac{1}{d}+k^2}{2} - \frac{d-1}{2d}\log\frac{1}{2d}
$$
$$
- \frac{\frac{1}{d}-k^2}{2}\log\frac{\frac{1}{d}-k^2}{2} - \frac{d-1}{2d}\log\frac{1}{2d}
$$
$$
= \frac{d-1}{d}\log(2d) - \frac{\frac{1}{d}+k^2}{2}\log\frac{\frac{1}{d}+k^2}{2}
$$
$$
- \frac{\frac{1}{d}-k^2}{2}\log\frac{\frac{1}{d}-k^2}{2}.
$$
(B17)

Now, note that one has

$$
k = \|F|\psi\rangle\| \leqslant \|F\|_\infty \leqslant \frac{1}{\sqrt{d}},
$$
(B18)

where the last inequality follows from Lemma 1. The expression (B17) is monotonically decreasing for k in the interval $[0, 1/\sqrt{d}]$. Hence, one has the lower bound

$$
H[\mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|)] \geqslant \frac{d-1}{d}\log(2d)
$$
$$
- \frac{\frac{1}{d}+\|F\|_\infty^2}{2}\log\frac{\frac{1}{d}+\|F\|_\infty^2}{2}
$$
$$
- \frac{\frac{1}{d}-\|F\|_\infty^2}{2}\log\frac{\frac{1}{d}-\|F\|_\infty^2}{2}.
$$
(B19)

Inserting this expression into Eq. (B14) with $\mathcal{E} = \mathcal{C}_{|+\rangle\langle+|,F}$, we then obtain Eq. (B13). ∎

Corollary 1. The Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ defined in Eq. (B3) is upper bounded as $\chi(\mathcal{C}_{\omega,F}) \leqslant 1/d$. In particular, for $d = 2$, one has the bound $\chi(\mathcal{C}_{\omega,F}) \leqslant 0.5$.

Proof. Immediate from the fact that the right-hand side of Eq. (B13) is monotonically decreasing with $\|F\|_\infty$, and that $\|F\|_\infty$ is upper bounded by $1/\sqrt{d}$ (Lemma 1). ∎

APPENDIX C: MAXIMIZATION OF THE HOLEVO INFORMATION FOR THE SUPERPOSITION OF INDEPENDENT DEPOLARIZING CHANNELS

Here we prove a series of results that enable a complete numerical maximization of the Holevo information of the channel (B3),

$$
\mathcal{C}_{\omega,F} : \rho \mapsto \frac{I/d + F\rho F^\dagger}{2} \otimes \omega + \frac{I/d - F\rho F^\dagger}{2} \otimes Z\omega Z,
$$
(C1)

over all input ensembles, over all states of the control system, and over all vacuum extensions of the completely depolarizing

channel. This Appendix makes use of notation introduced in the previous Appendices.

Let us start from the maximization over the vacuum extensions, which are in one-to-one correspondence with the possible operators F.

Lemma 3. Without loss of generality, the operator F that maximizes the Holevo information of the channel $\mathcal{C}_{\omega,F}$ can be taken to be of the form $F = a|0\rangle\langle0| + b|1\rangle\langle1|$, with $a^2 + b^2 \leqslant 1/d$, $a, b \geqslant 0$.

Proof. Using the singular value decomposition, F can be written as $F = UF'V$, where U and V are suitable unitary matrices, and F' is diagonal in the basis $\{|0\rangle, |1\rangle\}$. Now the capacity of the channel $\mathcal{C}_{\omega,F}$ is equal to the capacity of the channel $\mathcal{C}_{\omega,F'} = (\mathcal{U} \otimes \mathcal{I}_C)^\dagger \circ \mathcal{C}_{\omega,F} \circ \mathcal{V}^\dagger$, where $\mathcal{U}^\dagger$ and $\mathcal{V}^\dagger$ are the inverses of the unitary channels associated to the unitary matrices U and V, respectively, and $\mathcal{I}_C$ is the identity channel on the control system. Notice that F' is also a vacuum interference operator associated to the completely depolarizing channel. Hence, the maximization of the Holevo capacity can be restricted to channels with diagonal vacuum interference operator.

Next, we note that, for a vacuum extension of the completely depolarizing channel, the vacuum interference operator F must satisfy the condition $\mathrm{Tr}\, F^\dagger F \leqslant 1/d$ [31]. For an operator of the form $F = a|0\rangle\langle0| + b|1\rangle\langle1|$, this implies the inequality $|a|^2 + |b|^2 \leqslant 1/d$. Finally, we show that a, b can restricted to positive numbers. Let $W = a'|0\rangle\langle0| + b'|1\rangle\langle1|$, where $a' = \bar{a}/|a|$, $b' = \bar{b}/|b|$. Then $F'' := WF = FW = |a|\,|0\rangle\langle0| + |b|\,|1\rangle\langle1|$. The capacity of the channel $\mathcal{C}_{\omega,F''} = (\mathcal{W} \otimes \mathcal{I}_C) \circ \mathcal{C}_{\omega,F}$ (where $\mathcal{W}$ is the unitary channel associated with the unitary W) is equal to the capacity of the channel $\mathcal{C}_{\omega,F}$. Therefore, a maximization of the Holevo capacity can be restricted to vacuum interference operators with positive coefficients in the computational basis. ∎

Let us consider now the maximization over all possible ensembles. The key result here is that the maximization can be reduced to the optimization of d vectors with positive coefficients in the computational basis.

Lemma 4. When the operator F is diagonal in the computational basis, the input ensemble that maximizes the Holevo information after application of the channel $\mathcal{C}_{\omega,F}$ can be chosen without loss of generality to be of the form

$$
\left\{ \frac{p_x}{d},\, M^j |\psi_x\rangle\langle\psi_x| M^{j\dagger} \right\}_{x\in\{0,\dots,d-1\},\, j\in\{0,\dots,d-1\}},
$$
(C2)

where $(p_x)_{x\in\{0,\dots,d-1\}}$ is a probability distribution, M is the unitary operator $M := \sum_{m=0}^{d-1} \omega^m |m\rangle\langle m|$, $\omega := e^{2\pi i/d}$, and $|\psi_x\rangle$ is a unit vector with positive coefficients in the computational basis $\{|m\rangle\}_{m=0}^{d-1}$.

Proof. When F is diagonal, the channel $\mathcal{C}_{\omega,F}$ has the covariance property

$$
\mathcal{C}_{\omega,F} \circ \mathcal{U}_{\boldsymbol{\theta}} = (\mathcal{U}_{\boldsymbol{\theta}} \otimes \mathcal{I}_C) \circ \mathcal{C}_{\omega,F} \quad \forall \boldsymbol{\theta},
$$
(C3)

where $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_{d-1})$ is a vector of d phases, and $\mathcal{U}_{\boldsymbol{\theta}}$ is the unitary channel associated to the unitary matrix $U_{\boldsymbol{\theta}} = \sum_{m=0}^{d-1} e^{i\theta_m} |m\rangle\langle m|$. Note that, in particular, we have

$$
\mathcal{C}_{\omega,F} \circ \mathcal{M}^j = (\mathcal{M}^j \otimes \mathcal{I}_C) \circ \mathcal{C}_{\omega,F}, \quad j \in \{0, \dots, d-1\} \quad (C4)
$$

where $\mathcal{M}$ is the unitary channel associated to the unitary operator M defined in the statement of the lemma.

For covariant channels, Davies [63] showed that the optimal input ensembles can be chosen without loss of generality to be covariant. In our case, this means that the optimal ensemble can be chosen to be of the form

$$\mathsf{E} := \left\{ \frac{p_x}{d}, \, \mathcal{M}^j(\rho_x) \right\}_{x \in X, \, j \in \{0, \dots, d-1\}}, \tag{C5}$$

for some finite set X, some probability distribution $(p_x)_{x \in X}$, and some set of density matrices $(\rho_x)_{x \in X}$. In the same paper, Davies also showed that the ensemble can be chosen without loss of generality to consist of *pure* states, possibly at the price of increasing the size of the set X.

We now show that one can choose $|X| \leqslant d$ without loss of generality. Let E be an optimal covariant ensemble, and let

$$\langle \rho \rangle = \frac{1}{d} \sum_{j=0}^{d-1} \sum_{x \in X} p_x \, \mathcal{M}^j(\rho_x) \tag{C6}$$

be its average state. Fixing X, the set of covariant ensembles with average state $\langle \rho \rangle$ is a convex set. Since the Holevo information is a convex function of the ensemble [63], the maximization can be restricted without loss of generality to the extreme points.

Now, note that the covariant ensembles E are in one-to-one correspondence with covariant positive-operator-valued measures (POVMs) $(P_{x,j})_{x \in X, \, j \in \{0, \dots, d-1\}}$, via the correspondence

$$P_{x,j} := \frac{\mathcal{M}^j(\xi_x)}{d}, \quad \xi_x := \langle \rho \rangle^{-\frac{1}{2}} p_x \, \rho_x \, \langle \rho \rangle^{-\frac{1}{2}}. \tag{C7}$$

Since the correspondence is linear, the extreme ensembles are in one-to-one correspondence with the extreme POVMs. The latter have been characterized by one of us in Ref. [75], where it was shown that a necessary condition for extremality is that the ranks of the operators ξ_x, denoted by r_x, satisfy the condition

$$\sum_{x \in X} r_x^2 \leqslant \sum_{\mu} m_{\mu}^2, \tag{C8}$$

where the sum on the right-hand side runs over the irreducible representations (irreps) contained in the decomposition of the representation $\{M^j\}_{j=0}^{d-1}$, and m_{μ} is the multiplicity of the irrep μ. Now, the representation $\{M^j\}_{j=0}^{d-1}$ has d irreps, each with unit multiplicity. Hence, the bound becomes

$$\sum_{x \in X} r_x^2 \leqslant d. \tag{C9}$$

In particular, this means that the number of nonzero operators ξ_x is at most d.

In terms of the ensemble E, this means that the number of values of x with $p_x \neq 0$ is at most d. Hence, the maximization of the Holevo information can be restricted without loss of generality to covariant ensembles with $|X| \leqslant d$.

Recall that the optimal ensemble can be chosen without loss of generality to consist of pure states. The final step is to guarantee that these pure states have non-negative coefficients in the computational basis. For a covariant ensemble $\mathsf{E} = \{p_x/d, \, M^j|\psi_x\rangle\langle\psi_x|M^{j\dagger}\}$, let us expand each state as $|\psi_x\rangle = \sum_m |c_m| e^{i\theta_{x,m}} |m\rangle$, where $\{\theta_{x,m}\}$ are suitable phases. Then,

we can define the new states $|\psi_x'\rangle := U_{-\boldsymbol{\theta}_x}|\psi_x\rangle$, with $\boldsymbol{\theta}_x := (\theta_{x,0}, \dots, \theta_{x,d-1})$. By construction, these states have positive coefficients in the computational basis, and the corresponding ensemble $\mathsf{E}' := \{p_x/d, \, M^j|\psi_x'\rangle\langle\psi_x'|M^{j\dagger}\}$ gives rise to the same Holevo information as E, when fed into the channel $\mathcal{C}_{\omega,F}$. ∎

Corollary 2. When the operator F is diagonal in the computational basis, the Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ is given by

$$\chi(\mathcal{C}_{\omega,F}) = \max_{\{p_x, |\psi_x\rangle\}} \left\{ H\left[\mathcal{C}_{\omega,F}\left(\sum_{x,m} p_x \, |\langle m|\psi_x\rangle|^2 \, |m\rangle\langle m| \right) \right] \right.$$
$$\left. - \sum_x p_x \, H[\mathcal{C}_{\omega,F}(|\psi_x\rangle\langle\psi_x|)] \right\}, \tag{C10}$$

where the maximum is over the ensembles of d pure states with positive coefficients in the computational basis.

Proof. Immediate from the definition of the Holevo information for the ensemble obtained by applying channel $\mathcal{C}_{\omega,F}$ to the pure state ensemble in Lemma 4, using the relations

$$H[\mathcal{C}_{\omega,F}(M^j|\psi\rangle\langle\psi|M^{j\dagger})] = H[\mathcal{C}_{\omega,F}(|\psi\rangle\langle\psi|)], \tag{C11}$$

$$\frac{1}{d} \sum_{j=0}^{d-1} M^j|\psi\rangle\langle\psi|M^{j\dagger} = \sum_{m=0}^{d-1} |\langle m|\psi\rangle|^2 \, |m\rangle\langle m|, \tag{C12}$$

valid for every vector $|\psi\rangle$. ∎

For qubit messages ($d = 2$), we finally obtain an upper bound on the classical capacity:

Theorem 1. For every vacuum extension of the completely depolarizing channel and for every state of the control qubit, the classical capacity of the channel resulting from the superposition of two independent depolarizing qubit channels is upper bounded as

$$C(\mathcal{C}_{\omega,F}) \leqslant \max_{\substack{a \geqslant 0, \, b \geqslant 0 \\ a^2 + b^2 \leqslant 1/2}} \max_{0 \leqslant q, p_0, p_1 \leqslant 1} H[\mathcal{C}_{\omega,F}(\rho_q)]$$
$$- qH[\mathcal{C}_{\omega,F}(|\psi_0\rangle\langle\psi_0|)] - (1-q)H[\mathcal{C}_{\omega,F}(|\psi_1\rangle\langle\psi_1|)], \tag{C13}$$

$$|\psi_0\rangle = \sqrt{p_0}\,|0\rangle + \sqrt{1-p_0}\,|1\rangle,$$
$$|\psi_1\rangle = \sqrt{p_1}\,|1\rangle + \sqrt{1-p_1}\,|1\rangle,$$
$$\rho_q = [qp_0 + (1-q)p_1]|0\rangle\langle 0|$$
$$+ [q(1-p_0) + (1-q)(1-p_1)]|1\rangle\langle 1|. \tag{C14}$$

Proof. For $d = 2$, Proposition 1 guarantees that the channel $\mathcal{C}_{\omega,F}$ is entanglement breaking, and therefore its classical capacity is equal to the Holevo capacity. Lemma 2 guarantees that the maximum of the Holevo capacity is attained by the state $\omega = |+\rangle\langle +|$. Then, Lemma 4 guarantees the maximum of the Holevo capacity of the channel $\mathcal{C}_{|+\rangle\langle +|, F}$ can be obtained with a diagonal operator $F = a|0\rangle\langle 0| + b|1\rangle\langle 1|$, $a, b \geqslant 0$. The Holevo capacity of $\mathcal{C}_{|+\rangle\langle +|, F}$ can be computed explicitly using Corollary 2, with

$$|\psi_0\rangle := \sqrt{p_0}\,|0\rangle + \sqrt{1-p_0}\,|1\rangle$$
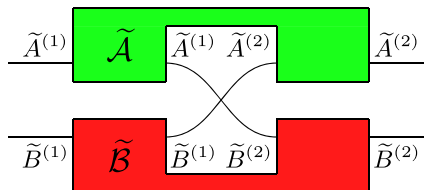$$|\psi_1\rangle := \sqrt{p_1}\,|1\rangle + \sqrt{1-p_1}\,|1\rangle. \tag{C15}$$

FIG. 10. The channel $\mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})$, obtained by connecting two vacuum-extended two-step channels $\widetilde{\mathcal{A}}$ (green) and $\widetilde{\mathcal{B}}$ (red), such that the output of the first use of each channel is connected to the input of the second use of the other channel.

Finally, an upper bound is obtained by relaxing the constraint on a and b to $a^2 + b^2 \leqslant 1/d$ (Lemma 4). ∎

APPENDIX D: TRANSMISSION OF A SINGLE PARTICLE THROUGH A NETWORK OF TWO-STEP CHANNELS

In the following we will use the notation introduced in Appendix A.

1. Derivation of Eq. (15) in the main text

Let $\mathcal{A}$ and $\mathcal{B}$ be two-step channels, with vacuum extensions $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{A}^{(1)}, \widetilde{A}^{(2)})$ and $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{B}^{(1)}, \widetilde{B}^{(2)})$. For simplicity, here we take all the systems $\widetilde{A}^{(1)}, \widetilde{A}^{(2)}, \widetilde{B}^{(1)}, \widetilde{B}^{(2)}$ to be isomorphic.

We now connect the two-step channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ in such a way that the output of the first use of each channel is fed into the input of the second use of the other channel, as in Fig. 10. This particular composition of two two-step channels is described by a supermap $\mathcal{Z}$ that maps pairs of channels in $\mathsf{Chan}(\widetilde{A}^{(1)}, \widetilde{A}^{(2)}) \times \mathsf{Chan}(\widetilde{B}^{(1)}, \widetilde{B}^{(2)})$ into bipartite channels in $\mathsf{Chan}(\widetilde{A}^{(1)} \otimes \widetilde{B}^{(1)} \to \widetilde{B}^{(2)} \otimes \widetilde{A}^{(2)})$.

We can now consider the scenario in which a single particle is sent in a superposition of going through the A port and the B port of the channel $\mathcal{Z}(\mathcal{R}_A, \mathcal{R}_B)$. Following Eq. (A1), the evolution of the particle is described by the superposition channel

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})] := \mathcal{U}^\dagger \circ \mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}) \circ \mathcal{U}, \qquad (D1)$$

with $\mathcal{U}$ defined as in Eqs. (A1) and (A2). The superposition channel $\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})]$ is illustrated in Fig. 11.

Let us apply the above construction to the special case where the channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are of the random unitary
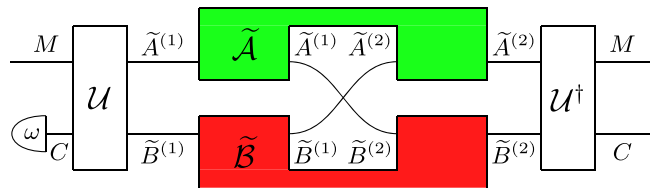

FIG. 11. The superposition channel $\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})]$ of two two-step channels $\mathcal{A}$ and $\mathcal{B}$, specified by the vacuum extensions $\widetilde{\mathcal{A}}$ (green) and $\widetilde{\mathcal{B}}$ (red), where the alternative paths traverse the two correlated channels in the opposite order. For the applications in this paper, the input of the control system C is fixed in the state ω, while the message system M is accessible to the sender.

form

$$\widetilde{\mathcal{A}} = \widetilde{\mathcal{R}}_A := \sum_{m,n} p_A(m, n)\, \widetilde{\mathcal{V}}_m^{(A)} \otimes \widetilde{\mathcal{V}}_n^{(A)},$$

$$\widetilde{\mathcal{B}} = \widetilde{\mathcal{R}}_B := \sum_{k,l} p_B(k, l)\, \widetilde{\mathcal{V}}_k^{(B)} \otimes \widetilde{\mathcal{V}}_l^{(B)}, \qquad (D2)$$

where $\widetilde{\mathcal{V}}_m^{(A)}$ and $\widetilde{\mathcal{V}}_k^{(B)}$ are the unitary channels corresponding to the unitary operators

$$\widetilde{V}_m^{(A)} := V_m^{(A)} \oplus e^{i\phi_m^{(A)}} |\mathrm{vac}\rangle\langle\mathrm{vac}|,$$

$$\widetilde{V}_k^{(B)} := V_k^{(B)} \oplus e^{i\phi_k^{(B)}} |\mathrm{vac}\rangle\langle\mathrm{vac}|, \qquad (D3)$$

respectively. With this choice, we have

$$\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)$$
$$= \sum_{m,n,k,l} p_A(m, n) p_B(k, l)\left(\widetilde{\mathcal{V}}_l^{(B)} \circ \widetilde{\mathcal{V}}_m^{(A)}\right) \otimes \left(\widetilde{\mathcal{V}}_n^{(A)} \circ \widetilde{\mathcal{V}}_k^{(B)}\right) \quad (D4)$$

and

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)](\cdot)$$
$$= \sum_{m,n,k,l} p_A(m, n)\, p_B(k, l)\, W_{mnkl}\, (\cdot)\, W_{mnkl}^\dagger, \qquad (D5)$$

with

$$W_{mnkl} := V_l^{(B)} V_m^{(A)}\, e^{i(\phi_k^{(B)} + \phi_n^{(A)})} \otimes |0\rangle\langle 0|$$
$$+ V_n^{(A)} V_k^{(B)}\, e^{i(\phi_m^{(A)} + \phi_l^{(B)})} \otimes |1\rangle\langle 1|. \qquad (D6)$$

This proves Eq. (15) in the main text.

2. Derivation of Eqs. (20) and (21) in the main text

For the control (in this case the path of the particle) initialized in the state ω, the superposition channel specified by the vacuum extension $\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)$ is given by

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)](\rho \otimes \omega) = \sum_{m,n,k,l} p_A(m, n)\, p_B(k, l)\, W_{mnkl}(\rho \otimes \omega) W_{mnkl}^\dagger$$

$$= \sum_{m,n,k,l} \left\{ p_A(m, n)\, p_B(k, l) V_l^{(B)} V_m^{(A)} \rho V_m^{(A)\dagger} V_l^{(B)\dagger} \otimes \omega_{00} |0\rangle\langle 0| \right.$$

$$+ p_A(m, n)\, p_B(k, l) V_n^{(A)} V_k^{(B)} \rho V_k^{(B)\dagger} V_n^{(A)\dagger} \otimes \omega_{11} |1\rangle\langle 1|$$

$$+ \left. p_A(m, n)\, p_B(k, l) V_l^{(B)} V_m^{(A)} \rho V_k^{(B)\dagger} V_n^{(A)\dagger} e^{i[\phi_n^{(A)} + \phi_k^{(B)} - \phi_m^{(A)} - \phi_l^{(B)}]} \otimes \omega_{01} |0\rangle\langle 1| + \mathrm{h.c.} \right\}$$

$$= \mathcal{R}_B \mathcal{R}_A(\rho) \otimes \omega_{00} |0\rangle\langle 0| + \mathcal{R}_A \mathcal{R}_B(\rho) \otimes \omega_{11} |1\rangle\langle 1| + \mathcal{K}(\rho) \otimes \omega_{01} |0\rangle\langle 1| + [\mathcal{K}(\rho)]^\dagger \otimes \omega_{10} |1\rangle\langle 0|,$$

where $\mathcal{K}$ is the linear map defined by

$$\mathcal{K}(\rho) := \sum_{m,n,k,l} p_A(m,n)\, p_B(k,l)\, V_l^{(B)} V_m^{(A)} \rho V_k^{(B)\dagger} V_n^{(A)\dagger}\, e^{i[\phi_n^{(A)}+\phi_k^{(B)}-\phi_m^{(A)}-\phi_l^{(B)}]}. \tag{D7}$$

We now restrict our attention to the case where

(1) the two channels $\widetilde{\mathcal{R}}_A$ and $\widetilde{\mathcal{R}}_B$ are identical [this implies that one can choose without loss of generality $p_A(m,n) = p_B(m,n) =: p(m,n)$ for every m and n, $V_m^{(A)} = V_m^{(B)} =: V_m$, and $\phi_m^{(A)} = \phi_m^{(B)} =: \phi_m$ for every m];

(2) the probability distribution $p(m,n)$ is symmetric, namely, $p(m,n) = p(n,m)$ for every m, n.

Under these conditions, the operator $\mathcal{K}(\rho)$ is self-adjoint for every density matrix ρ, and the effective channel can be rewritten as

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)](\rho \otimes \omega) = \frac{\mathcal{R}^2(\rho) + \mathcal{K}(\rho)}{2} \otimes \omega$$
$$+ \frac{\mathcal{R}^2(\rho) - \mathcal{K}(\rho)}{2} \otimes Z\omega Z, \tag{D8}$$

with

$$\mathcal{R}(\rho) := \sum_{m,n} p(m,n)\, V_m \rho V_m^{\dagger}. \tag{D9}$$

In particular, suppose that the unitaries $\{V_m\}_{m=0}^{d^2-1}$ form an orthogonal basis, and that the probability $p(m,n)$ has the form $p(m,n) = \delta_{n,\sigma(m)}/d^2$, for a permutation σ that makes $p(m,n)$ symmetric. In this case, Eq. (D8) becomes

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B)](\rho \otimes \omega) = \frac{I/d + \mathcal{K}(\rho)}{2} \otimes \omega$$
$$+ \frac{I/d - \mathcal{K}(\rho)}{2} \otimes Z\omega Z, \tag{D10}$$

with

$$\mathcal{K}(\rho) = \frac{1}{d^4} \sum_{m,k} V_{\sigma(k)}^{(B)} V_m^{(A)} \rho V_k^{(B)\dagger} V_{\sigma(m)}^{(A)\dagger}\, e^{i[\phi_{\sigma(m)}^{(A)}+\phi_k^{(B)}-\phi_m^{(A)}-\phi_{\sigma(k)}^{(B)}]}. \tag{D11}$$

Setting $d = 2$ and choosing σ to be the permutation that exchanges 0 with 1, and 2 with 3, we obtain Eqs. (20) and (21) of the main text.

APPENDIX E: PROOFS OF THE STATEMENTS IN SEC. IV B

Here we consider the scenario of Fig. 11, in the special case where the two-step channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are of the product form $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 \otimes \widetilde{\mathcal{A}}_2$ and $\widetilde{\mathcal{B}} = \widetilde{\mathcal{B}}_1 \otimes \widetilde{\mathcal{B}}_2$, respectively. In this case, the combination of the channels in the network of Fig. 10 gives the bipartite channel

$$\mathcal{Z}(\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{B}}) = \widetilde{\mathcal{B}}_2 \widetilde{\mathcal{A}}_1 \otimes \widetilde{\mathcal{A}}_2 \widetilde{\mathcal{B}}_1. \tag{E1}$$

When a single particle is sent into one of the two ports of this channel, the resulting evolution is described by the superposition channel

$$\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{B}})] = \mathcal{S}(\widetilde{\mathcal{B}}_2 \widetilde{\mathcal{A}}_1 \otimes \widetilde{\mathcal{A}}_2 \widetilde{\mathcal{B}}_1), \tag{E2}$$

where $\mathcal{S}$ is the supermap defined in Eq. (A1).

We now restrict our attention to the case where the channels $\widetilde{\mathcal{A}}_1$, $\widetilde{\mathcal{A}}_2$, $\widetilde{\mathcal{B}}_1$, and $\widetilde{\mathcal{B}}_2$ are all equal to each other, and are all equal to $\widetilde{\mathcal{D}}$, a vacuum extension of the completely depolarizing channel. In this case, the action of the superposition channel on a generic product state $\rho \otimes \omega$ is

$$\mathcal{S}(\widetilde{\mathcal{D}}^2 \otimes \widetilde{\mathcal{D}}^2)(\rho \otimes \omega) \tag{E3}$$
$$= \frac{I/d + F^2 \rho F^{2\dagger}}{2} \otimes \omega + \frac{I/d - F^2 \rho F^{2\dagger}}{2} \otimes Z\omega Z, \tag{E4}$$

where F is the vacuum interference operator associated to channel $\widetilde{\mathcal{D}}$. The above equation follows from Eq. (B3) and from the observation that the vacuum interference operator of $\widetilde{\mathcal{D}}^2$ is F^2.

Note that one has the equality

$$\mathcal{S}(\widetilde{\mathcal{D}}^2 \otimes \widetilde{\mathcal{D}}^2)(\rho \otimes \omega) \equiv \mathcal{C}_{\omega, F^2}(\rho), \tag{E5}$$

using the notation of Eq. (B3). That is, in the lack of correlations the configuration of channels depicted in Fig. 11 gives rise to the effective channel in Eq. (B1), with F replaced by F^2. This means that all of the results in Appendixes B and C apply to this scenario as well, with F replaced by F^2. In particular, the classical capacity can be determined numerically using Theorem 1, with the maximization constraint now being that for the vacuum interference operator $F^2 = g\,|0\rangle\langle 0| + h\,|1\rangle\langle 1|$, $g + h \leqslant 1/d$, where $g, h \geqslant 0$.

The classical capacity of the channels $\mathcal{C}_{\omega, F}$ and $\mathcal{C}_{\omega, F^2}$ can be evaluated numerically. For the cases where each completely
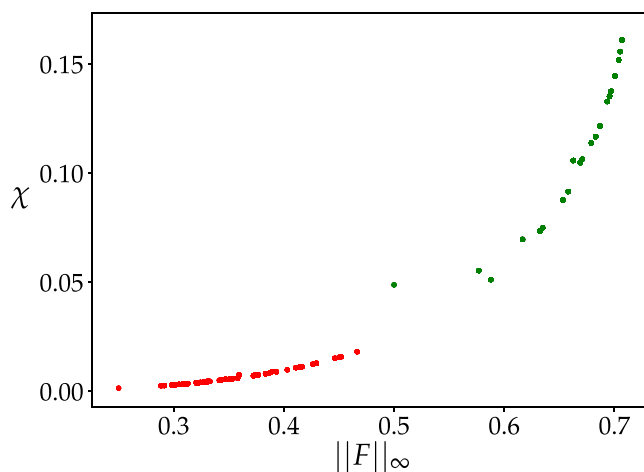

FIG. 12. Green: A plot of the classical capacity against $||F||_{\infty}$ for the channel $\mathcal{C}_{\omega, F}$. Red: A plot of the classical capacity against $||F||_{\infty}^2$ for the channel $\mathcal{C}_{\omega, F^2}$. In both cases $F = \sum_{m=0}^{3} \frac{1}{4} e^{-i\phi_m} V_m$ and is sampled over the phase parameters $\{\phi_1, \phi_2, \phi_3\}$ with a numerical precision of $\pi/8$ for each parameter. We set $\phi_0 = 0$ without loss of generality, as $F\rho F^{\dagger}$ is invariant under the phase group U(1). The classical capacity is here equal to the Holevo capacity (see Appendix B) and the Holevo capacity was calculated using the methods outlined in Appendix C.

depolarizing channel is implemented by a random unitary channel [cf. Eqs. (4) and (15), respectively, in the main text], Fig. 12 show a scatter plot with the capacities of both channels in the same graph against the norm of the corresponding vacuum interference operator, F or F^2, for same combination of phases ϕ_1, ϕ_2, ϕ_3 as shown in Figs. 4 and 6.

[1] C. H. Bennett and G. Brassard, Quantum cryptography: public key distribution and coin tossing, in *International Conference on Computers, Systems and Signal Processing (Bangalore, India, 1984)* (IEEE Computer Society, Washington, DC, 1984), pp. 175–9.

[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[3] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).

[4] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation, in *Quantum Information Science and its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics* (American Mathematical Society, Providence, RI, 2010), Vol. 68, pp. 13–58.

[5] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, 2013).

[6] C. Macchiavello and G. M. Palma, Entanglement-enhanced information transmission over a quantum channel with correlated noise, Phys. Rev. A **65**, 050301(R) (2002).

[7] D. Kretschmann and R. F. Werner, Quantum channels with memory, Phys. Rev. A **72**, 062323 (2005).

[8] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Quantum channels and memory effects, Rev. Mod. Phys. **86**, 1203 (2014).

[9] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, Phys. Rev. A **97**, 012127 (2018).

[10] J. L. Ball and K. Banaszek, Hybrid noiseless subsystems for quantum communication over optical fibres, J. Phys. A: Math. Gen. **39**, L1 (2005).

[11] C. Bonato, M. Aspelmeyer, T. Jennewein, C. Pernechele, P. Villoresi, and A. Zeilinger, Influence of satellite motion on polarization qubits in a space-earth quantum communication link, Opt. Express **14**, 10050 (2006).

[12] G. Chiribella, M. Dall'Arno, G. M. D'Ariano, C. Macchiavello, and P. Perinotti, Quantum error correction with degenerate codes for correlated noise, Phys. Rev. A **83**, 052305 (2011).

[13] V. Giovannetti and R. Fazio, Information-capacity description of spin-chain correlations, Phys. Rev. A **71**, 032314 (2005).

[14] C. Macchiavello, G. M. Palma, and S. Virmani, Transition behavior in the channel capacity of two-qubit channels with memory, Phys. Rev. A **69**, 010303(R) (2004).

[15] G. Ruggeri, G. Soliani, V. Giovannetti, and S. Mancini, Information transmission through lossy bosonic memory channels, Europhys. Lett. **70**, 719 (2005).

[16] N. J. Cerf, J. Clavareau, C. Macchiavello, and J. Roland, Quantum entanglement enhances the capacity of bosonic channels with memory, Phys. Rev. A **72**, 042330 (2005).

[17] V. Giovannetti and S. Mancini, Bosonic memory channels, Phys. Rev. A **71**, 062304 (2005).

[18] J. Ball, A. Dragan, and K. Banaszek, Exploiting entanglement in communication channels with correlated noise, Phys. Rev. A **69**, 042324 (2004).

[19] K. Banaszek, A. Dragan, W. Wasilewski, and C. Radzewicz, Experimental Demonstration of Entanglement-Enhanced Classical Communication Over a Quantum Channel with Correlated Noise, Phys. Rev. Lett. **92**, 257901 (2004).

[20] G. Bowen and S. Mancini, Quantum channels with a finite memory, Phys. Rev. A **69**, 012306 (2004).

[21] M. B. Plenio and S. Virmani, Spin Chains and Channels with Memory, Phys. Rev. Lett. **99**, 120504 (2007).

[22] A. Bayat, D. Burgarth, S. Mancini, and S. Bose, Memory effects in spin-chain channels for information transmission, Phys. Rev. A **77**, 050306(R) (2008).

[23] E. Karpov, D. Daems, and N. J. Cerf, Entanglement-enhanced classical capacity of quantum communication channels with memory in arbitrary dimensions, Phys. Rev. A **74**, 032320 (2006).

[24] L. Memarzadeh, C. Macchiavello, and S. Mancini, Recovering quantum information through partial access to the environment, New J. Phys. **13**, 103031 (2011).

[25] X. Xiao, Y. Yao, Y.-M. Xie, X.-H. Wang, and Y.-L. Li, Protecting entanglement from correlated amplitude damping channel using weak measurement and quantum measurement reversal, Quantum Inf. Process. **15**, 3881 (2016).

[26] A. D'Arrigo, G. Benenti, and G. Falci, Transmission of classical and quantum information through a quantum memory channel with damping, Eur. Phys. J. D **66**, 147 (2012).

[27] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, Superpositions of Time Evolutions of a Quantum System and a Quantum Time-Translation Machine, Phys. Rev. Lett. **64**, 2965 (1990).

[28] D. K. L. Oi, Interference of Quantum Channels, Phys. Rev. Lett. **91**, 067902 (2003).

[29] J. Åberg, Subspace preservation, subspace locality, and gluing of completely positive maps, Ann. Phys. **313**, 326 (2004).

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

[31] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, Communication through coherent control of quantum channels, Quantum **4**, 333 (2020).

[32] G. Chiribella and H. Kristjánsson, Quantum Shannon theory with superpositions of trajectories, Proc. R. Soc. A **475**, 20180903 (2019).

[33] Q. Dong, S. Nakayama, A. Soeda, and M. Murao, Controlled quantum operations and combs, and their applications to universal controllization of divisible unitary operations, arXiv:1911.01645.

[34] A. Vanrietvelde and G. Chiribella, Universal control of quantum processes using sector-preserving channels, arXiv:2106.12463.

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

[36] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Beyond quantum computers, arXiv:0912.0195v1.

[37] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).

[38] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. 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).

[39] 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).

[40] 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).

[41] Yu 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).

[42] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Increasing communication capacity via superposition of order, Phys. Rev. Research **2**, 033292 (2020).

[43] K. Goswami and J. Romero, Experiments on quantum causality, AVS Quantum Sci. **2**, 037101 (2020).

[44] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. A. Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella *et al.*, Experimental quantum communication enhancement by superposing trajectories, Phys. Rev. Research **3**, 013093 (2021).

[45] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, Phys. Rev. Lett. **120**, 120502 (2018).

[46] S. Salek, D. Ebler, and G. Chiribella, Quantum communication in a superposition of causal orders, arXiv:1809.06655.

[47] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, Indefinite causal order enables perfect quantum communication with zero capacity channels, New J. Phys. **23**, 033039 (2021).

[48] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, Communication enhancement through quantum coherent control of n channels in an indefinite causal-order scenario, Entropy **21**, 1012 (2019).

[49] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, Sending classical information via three noisy channels in superposition of causal orders, Phys. Rev. A **101**, 012346 (2020).

[50] G. Chiribella, M. Wilson, and H.-F. Chau, Quantum and Classical Data Transmission Through Completely Depolarising Channels in a Superposition of Cyclic Orders, Phys. Rev. Lett. **127**, 190502 (2021).

[51] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, Phys. Rev. Lett. **101**, 060401 (2008).

[52] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[53] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, Colloquium: Non-markovian dynamics in open quantum systems, Rev. Mod. Phys. **88**, 021002 (2016).

[54] P. C. Humphreys, B. J. Metcalf, J. B. Spring, M. Moore, X.-M. Jin, M. Barbieri, W. S. Kolthammer, and I. A. Walmsley, Linear Optical Quantum Computing in a Single Spatial Mode, Phys. Rev. Lett. **111**, 150501 (2013).

[55] J. M. Donohue, M. Agnew, J. Lavoie, and K. J. Resch, Coherent Ultrafast Measurement of Time-Bin Encoded Photons, Phys. Rev. Lett. **111**, 153602 (2013).

[56] X.-H. Li and S. Ghose, Hyperentanglement concentration for time-bin and polarization hyperentangled photons, Phys. Rev. A **91**, 062302 (2015).

[57] J. M. Donohue, J. Lavoie, and K. J. Resch, Ultrafast Time-Division Demultiplexing of Polarization-Entangled Photons, Phys. Rev. Lett. **113**, 163602 (2014).

[58] F. Ghafari, N. Tischler, C. Di Franco, J. Thompson, M. Gu, and G. J. Pryde, Interfering trajectories in experimental quantum-enhanced stochastic simulation, Nat. Commun. **10**, 1630 (2019).

[59] I. Marcikic, H. de Riedmatten, W. Tittel, V. Scarani, H. Zbinden, and N. Gisin, Time-bin entangled qubits for quantum communication created by femtosecond pulses, Phys. Rev. A **66**, 062308 (2002).

[60] G. Birkhoff, Three observations on linear algebra, Univ. Nac. Tacuman, Rev. Ser. A **5**, 147 (1946).

[61] M. Horodecki, P. W. Shor, and M. B. Ruskai, Entanglement breaking channels, Rev. Math. Phys. **15**, 629 (2003).

[62] P. W. Shor, Additivity of the classical capacity of entanglement-breaking quantum channels, J. Math. Phys. **43**, 4334 (2002).

[63] E. Davies, Information and quantum measurement, IEEE Trans. Inf. Theory **24**, 596 (1978).

[64] C. A. Fuchs, Nonorthogonal Quantum States Maximize Classical Information Capacity, Phys. Rev. Lett. **79**, 1162 (1997).

[65] C. King, M. Nathanson, and M. B. Ruskai, Qubit Channels Can Require More Than Two Inputs to Achieve Capacity, Phys. Rev. Lett. **88**, 057901 (2002).

[66] M. Hayashi, H. Imai, K. Matsumoto, M. B. Ruskai, and T. Shimono, Qubit channels which require four inputs to achieve capacity: Implications for additivity conjectures, Quantum Inf. Comput. **5**, 13 (2005).

[67] H. Kristjánsson, Y. Zhong, A. Munson, and G. Chiribella (unpublished).

[68] F. Del Santo and B. Dakić, Two-Way Communication with a Single Quantum Particle, Phys. Rev. Lett. **120**, 060503 (2018).

[69] J. Åberg, Operations and single-particle interferometry, Phys. Rev. A **70**, 012103 (2004).

[70] 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).

[71] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: quantum supermaps, Europhys. Lett. **83**, 30004 (2008).

[72] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[73] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of n-particle mixed states: Necessary and sufficient conditions in terms of linear maps, Phys. Lett. A **283**, 1 (2001).

[74] A. S. Holevo, Remarks on the classical capacity of quantum channel, arXiv:quant-ph/0212025.

[75] G. Chiribella and G. M. D'Ariano, Extremal covariant measurements, J. Math. Phys. **47**, 092107 (2006).